Electromagnetic Field with Constraints and Papapetrou Equation

Zafar Y. Turakulov[a,b] and Alisher T. Muminov[c]

[a] National University of Uzbekistan, Vuzgorodok, Tashkent 700174, Uzbekistan
[b] Astronomy Institute, and Institute of Nuclear Physics, Ulughbek, Tashkent 702132, Uzbekistan
[c] Institute of Applied Physics affiliated to the National University of Uzbekistan, Tashkent, Uzbekistan

Reprint requests to Dr. Z. Y. T.; E-mail: tzafar@astrin.uzsci.net

Z. Naturforsch. **61a,** 146 – 152 (2006); received January 12, 2006

It is shown that a geometric optical description of the electromagnetic wave with respect to its polarization in a curved space-time can be obtained straightforwardly from the classical variational principle for the electromagnetic field. For this purpose the entire functional space of electromagnetic fields must be reduced to its subspace of locally plane monochromatic waves. We have formulated the constraints under which this can be achieved. These constraints introduce variables of another kind which specify a field of local frames associated with the wave. They contain some congruence with null-curves. The Lagrangian for constrained electromagnetic fields contains variables of two kinds, namely a congruence of null-curves and the field itself. This in turn yields two kinds of Euler-Lagrange equations. The equations of the first kind are trivial due to the constraints imposed. The variation of the curves yields the Papapetrou equations for a classical massless particle with helicity 1.

Key words: General Relativity; Electromagnetic Waves; Circular Polarization; Papapetrou Equation.

1. Introduction

Quantum mechanics provides an exhaustive description of motion of a particle, if the typical length of the run is comparable with the wavelength in atomic and sub-atomic scales. When considering the motion of a particle on scales which are apparently non-comparable with the typical wavelength, that quantum mechanical picture becomes less convenient. To use it, one passes to an asymptotical behavior of incident and scattered waves. On astrophysical scales, particularly, when studying the deflection of light in a gravitational field, classical mechanics is evidently more convenient and one prefers to consider photons as classical massless particles moving along a null-geodesic in space-time. This approach to light propagation in curved space-time is quite satisfactory if the photon can be considered as a scalar particle. If, however, its polarization becomes important the question arises, how to include this into the classical mechanical description.

If the wavelength is large enough, i.e., if the photon momentum is relatively small, the common classical mechanical description becomes incomplete. The reason is that under some conditions the spin of the particle, which usually is regarded as zero, becomes comparable with some components of its orbital momentum. This may happen, for example, for a light beam incident to a Schwarzschild black hole. Due to the classical mechanical considerations each photon of the beam has zero longitudinal component of orbital momentum. However, its spin is also longitudinal and, hence, contributes to the total longitudinal component of angular momentum [1]. Since, on the one hand, spin is always collinear to the momentum of the photon, and, on the other hand, gravitational deflection of the beam changes the momentum, this description contradicts the conservation law of total angular momentum. It is clear that in order to have a correct picture one should take the spin of the photon into account such that the total angular momentum is conserved. If it is done, the sum of the longitudinal components of the orbital and the inner angular momenta is constant. Thus the first one is not zero after deflecting. If the photon momentum is small enough, this effect can change the shape of the scattered beam. Thus, there exist situations in which the well-known corpuscular theory does not work.

While in case of spinless particles one has a choice between the usual (corpuscular) and wave mechanics depending on the to scales under consideration, in case of particles with helicity the former one does not apply, so, the only possibility is to use the latter one. The provides a description in terms of special functions, say Legendre polynomials, which are convenient, if their label is not very high. However, in some real situations the order of the polynomials are of the order of astronomical distances measured in wavelength units. In these situations a corpuscular description in which the spin of the particle is taken into account properly, would be much more convenient. The goal of the present work is to construct a model of the photon with given momentum and helicity ± 1 in a curved space-time.

2. Statement of Problem

The desired model is assumed to provide certain world lines of a massless particle which has helicity 1 and, at the same time, describes on electromagnetic field which everywhere represents a locally plane wave. The main difficulty is that the spin of the particle is quantized, thus, the desired construction must contain both classical and quantum degrees of freedom. The notion of a classical particle with a quantum spin seems to be one of the simplest systems of mixed classical-quantum nature, and attempts to built a correct theory of this object are lasting for decades [2]. Our aim is somewhat broader, because we not only try to obtain an equation for the particle world line, but also a locally plane wave, in other words, to describe the propagation of circularly polarized photons in terms of both geometric and wave optics.

An attempt to build such a model was made in [3, 4] motivated by a problem of light propagation in Schwarzschild space-time. The presence of an electromagnetic field raises new questions, namely how to combine the field in the space-time and the helicity which must be attached to the still unknown world line. In [3] the electromagnetic field was removed by a vector field which was defined on the world lines. This substitution made it possible to combine the wave and the world line under the assumption that if the lines are found properly then the fields attached to them constitute the entire electromagnetic field in the space-time.

In the present work we revise this approach. Instead of specifying a functional space of curves and attaching a vector field to each curve, we assume that the same result can be obtained from the pure electromagnetic Lagrangian. The main idea of this work is that, after all, both geometric and wave optics should follow from pure electromagnetic theory, therefore, the model should be built in the framework of the theory of electromagnetic fields. We start with the well-known form of the Lagrangian of electromagnetic field and restrict the functional space of the field variables to its subspace of fields behaving as locally plane waves. Since restrictions of this sort are known as constraints we assume that putting relevant constraints on the Lagrangian leads to special Euler-Lagrange equations for the waves in question, and the desired description containing both geometric and wave optics follows from it.

3. Locally Plane Wave and Associated Orthonormal Frame

The notion of plane wave in space-time differs from that in space where the wave vector is orthogonal to the hyperplanes, and cannot be tangent to them. In space-time the wave vector is orthogonal to the hyperplanes and, at the same time, tangent to them. To see this consider flat space-time and Cartesian coordinates $\{t,x,y,z\}$ in it which are chosen in such a way that the wave has the phase $\phi = \omega(t-z)$. The phase takes constant values on luminal hyperplanes $t = z$ and the wave vector $\boldsymbol{e}_- = 2^{-1/2}(\partial_t + \partial_z)$ is tangent to the hyperplanes:

$$\boldsymbol{e}_- \circ \phi = 0. \qquad (1)$$

At the same time the wave vector is orthogonal to the hyperplanes because, due to pseudo-Euclidean metric, any null-vector is orthogonal to itself. The wave propagates along this vector, therefore $\boldsymbol{e}_-$ must be identified with the vector of velocity and plays the role of the velocity of the photon in a corpuscular model. Now we use these two objects to construct an orthonormal frame associated with the wave.

By construction, there exists an isotropic vector $\boldsymbol{e}_-$ tangent everywhere to the surfaces of constant phase, and, therefore, orthogonal to them. We introduce one more isotropic vector $\boldsymbol{e}_+$ whose direction is arbitrary, requiring only that its scalar product with the vector $\boldsymbol{e}_-$ is equal to one. As this is done we can introduce two unit space-like vectors $\boldsymbol{e}_\alpha$ which are orthogonal to each other and to $\boldsymbol{e}_\pm$. The four vectors defined this way constitute a local orthonormal frame.

In Minkowski space-time and Cartesian coordinates considered above the space-like vectors $\boldsymbol{e}_\alpha$ $(\alpha = 1,2)$ are defined as follows:

$$\boldsymbol{e}_1 = \partial_x, \quad \boldsymbol{e}_2 = \partial_y.$$

They are also tangent to the wave fronts and orthogonal to them. These vectors are used for specifying the polarization of the wave. The four vectors $\boldsymbol{e}_\alpha, \boldsymbol{e}_\pm$, where the coordinates are chosen such that $\boldsymbol{e}_+ = 2^{-1/2}$ $(\partial_t - \partial_z)$, form an orthonormal frame with the metric

$$\begin{aligned} &\langle \boldsymbol{e}_+, \boldsymbol{e}_+\rangle = \langle \boldsymbol{e}_-, \boldsymbol{e}_-\rangle = \langle \boldsymbol{e}_\pm, \boldsymbol{e}_\alpha\rangle = 0, \\ &\langle \boldsymbol{e}_-, \boldsymbol{e}_+\rangle = 1,\ \langle \boldsymbol{e}_\alpha, \boldsymbol{e}_\beta\rangle = -\delta_{\alpha\beta},\ \alpha, \beta = 1,2. \end{aligned} \qquad (2)$$

The frame of 1-forms dual to it is

$$\begin{aligned} &\theta^- = \mathrm{d}t + \mathrm{d}z, \quad \theta^+ = \mathrm{d}t - \mathrm{d}z, \\ &\theta^1 = \mathrm{d}x, \quad \theta^2 = \mathrm{d}y. \end{aligned}$$

Now, let us pass to a curved space-time and account for the main features of an electromagnetic wave which can be called locally plane and monochromatic. We start with a wave which locally possesses a scalar function ϕ called its phase. The gradient of this function is an isotropic 1-form and its level (hyper-)surfaces are wave fronts, i.e., it is possible to introduce local Cartesian coordinates in which the wave can be represented locally in the way just discussed. This is possible under some special condition, due to which the wavelength is much less than typical scales specified by the space-time curvature. Hereafter we assume that this condition is satisfied. The field of orthonormal frames associated with the wave can be constructed similarly.

After this is done it remains to fix one detail. The point is that the metric of this frame, given by (2), remains unchanged if one of the isotropic vectors is multiplied by an arbitrary factor and another one is divided by it. Employing this operation we can fix the action of the vector $\boldsymbol{e}_-$ on the phase:

$$\boldsymbol{e}_+ \circ \phi = \omega, \quad \omega = \text{const.}, \qquad (3)$$

while, by construction, (1) remains valid. Now, as the orthonormal vector frame associated with the wave $\{\boldsymbol{e}_\pm, \boldsymbol{e}_\alpha\}$ is defined, one can find out the orthonormal covector frame $\{\theta^\pm, \theta^\alpha\}$ dual to it and the connecting 1-form $\omega_{a\cdot}^{\ b} \equiv \gamma_{ca\cdot}^{\ \ b}\, \theta^c$ for the frame. Hereafter we assume that this is done, and pass to consider the constraints imposed on electromagnetic field to leave only locally plane waves.

4. Electromagnetic Field with Constraints

If the field in question carries everywhere a locally plane wave, there exists a congruence of null-curves $\{x^i(s)\}$ which are integral lines of the vector $\boldsymbol{e}_-$. Since this vector is identified with the velocity, the wave vector is colinear to it and the potential of the field α has no $\pm$ components. Thus, we have a constraint given by

$$\alpha = A_\beta\, \theta^\beta, \quad \beta = 1,2. \qquad (4)$$

The vectors $\boldsymbol{e}_1, \boldsymbol{e}_2$ are chosen to specify the polarization of the wave, hence, their span represents the local (two-dimensional space-like) wave front. Therefore we can neglect changes of the potential in their directions. This constraint can be written as

$$\boldsymbol{e}_1 \circ \alpha = \boldsymbol{e}_2 \circ \alpha = 0, \qquad (5)$$

where the operator $\boldsymbol{e}_\beta \circ$ stands for differentiation along the vector $\boldsymbol{e}_\beta$.

By analogy with plane waves in Cartesian coordinates the amplitudes can be represented in the form

$$A_\beta = a_\beta \mathrm{e}^{\mathrm{i}\phi}, \qquad (6)$$

where the contants a_β are complex numbers chosen such that the wave has left or right circular polarization and the function ϕ specifies the phase of the wave. In these denotions the constraint (5) coincides with (1). Finally, all the constraints imposed above reduce to the following: the 1-form of field potential has only one non-zero derivative

$$\boldsymbol{e}_+ \circ \alpha = \mathrm{i}\omega\, \alpha, \quad \boldsymbol{e}_- \circ \alpha = \boldsymbol{e}_\beta \circ \alpha = 0, \qquad (7)$$

where ϕ and ω do not depend on the parameter s.

5. Action Principle for Constrained Fields

The action functional for the electromagnetic field is given by the well-known integral

$$\mathcal{A} = \int \mathrm{d}\alpha \wedge {}^*\mathrm{d}\alpha = \int \langle \mathrm{d}\alpha, \mathrm{d}\alpha\rangle \varepsilon, \qquad (8)$$

where α is the 1-form of the potential of the field, $\langle , \rangle$ stands for the scalar product and ε denotes the unit 4-form: $\varepsilon \equiv (\varepsilon_{ijkl}/4!)\ \theta^i \wedge \theta^j \wedge \theta^k \wedge \theta^l$ that corresponds to the four-dimensional integration in the space-time. Straightforward computation of the variation of the action (8) yields Maxwell's equations, valid for all possible shapes of the electromagnetic field. Our goal is to restrict the functional space of the field to the subspace of fields which carry everywhere locally plane waves by imposing the constraints (1, 3 – 7).

For this purpose we expand the quadric form $< \mathrm{d}\alpha, \mathrm{d}\alpha >$ in the frame $\{\theta\}$ as follows:

$$\begin{aligned}
&\langle \mathrm{d}\alpha, \mathrm{d}\alpha \rangle = \\
&(\boldsymbol{e}_a \circ A_b)(\boldsymbol{e}_c \circ A_d) \left\langle \theta^a \wedge \theta^b, \theta^c \wedge \theta^d \right\rangle = \\
&\sum_{a,c=\pm} (\boldsymbol{e}_a \circ A_b)(\boldsymbol{e}_c \circ A_d) \left\langle \theta^a \wedge \theta^b, \theta^c \wedge \theta^d \right\rangle = \\
&\sum_{a,c=\pm} (\boldsymbol{e}_a \circ A_b)(\boldsymbol{e}_c \circ A_d) \langle \theta^a, \theta^c \rangle \left\langle \theta^b, \theta^d \right\rangle = \\
&\sum_{a,c=\pm} \langle \boldsymbol{e}_a \circ \boldsymbol{A}, \boldsymbol{e}_c \circ \boldsymbol{A} \rangle \langle \theta^a, \theta^c \rangle, \\
&\text{where} \quad \boldsymbol{A} = A^\beta \boldsymbol{e}_\beta.
\end{aligned}$$

Though most of terms of the expansion are zero, some of them have non-zero variations. Substituting the constrained Lagrangian into the action integral (8) yields:

$$\begin{aligned}
\mathcal{A} = \frac{1}{2} \int \{ &\langle \boldsymbol{e}_+ \circ \bar{\boldsymbol{A}}, \boldsymbol{e}_+ \circ \boldsymbol{A} \rangle \langle \theta^+, \theta^+ \rangle \\
&+ \langle \boldsymbol{e}_- \circ \bar{\boldsymbol{A}}, \boldsymbol{e}_- \circ \boldsymbol{A} \rangle \langle \theta^-, \theta^- \rangle \\
&+ (\langle \boldsymbol{e}_- \circ \bar{\boldsymbol{A}}, \boldsymbol{e}_+ \circ \boldsymbol{A} \rangle + \text{C. C.}) \langle \theta^-, \theta^+ \rangle \} \, \varepsilon,
\end{aligned} \tag{9}$$

where we take into account the fact that for convenience we use complex valued field components. As usual, the components $\boldsymbol{A}$ and their complex conjugates $\bar{\boldsymbol{A}}$ are regarded as independent variables. Due to the constraints $(1, 3-7)$ the Lagrangian under the integral (9) can be transformed as follows:

$$\begin{aligned}
&\{ \langle \boldsymbol{e}_+ \circ \bar{\boldsymbol{A}}, \boldsymbol{e}_+ \circ \boldsymbol{A} \rangle \langle \theta^+, \theta^+ \rangle \\
&\quad + \langle \boldsymbol{e}_- \circ \bar{\boldsymbol{A}}, \boldsymbol{e}_- \circ \boldsymbol{A} \rangle \langle \theta^-, \theta^- \rangle \\
&\quad + (\langle \boldsymbol{e}_- \circ \bar{\boldsymbol{A}}, \boldsymbol{e}_+ \circ \boldsymbol{A} \rangle + \text{C. C.}) \langle \theta^-, \theta^+ \rangle \} \\
&\quad = \omega^2 \langle \bar{\boldsymbol{A}}, \boldsymbol{A} \rangle \langle \theta^+, \theta^+ \rangle + \mathrm{i}\omega \left\langle \boldsymbol{A}, \dot{\bar{\boldsymbol{A}}} \right\rangle + \text{C. C.} \,.
\end{aligned}$$

The third term can be ignored because, as will be shown below, the action of the vector $\boldsymbol{e}_-$ annulates the field, consequently, the expression $\langle \boldsymbol{e}_- \circ \bar{\boldsymbol{A}}, \boldsymbol{e}_- \circ \boldsymbol{A} \rangle$ is the product of two zero factors, hence, both the third term itself and its variation are identically zero. Though, due to the constraint (1) the term $\dot{\bar{\boldsymbol{A}}}$ (and $\dot{\boldsymbol{A}}$) is equal to zero, we do not ignore it because its variation plays an important role in the action principle. Thus, finally, the action integral has the form

$$\begin{aligned}
&\mathcal{A} = \int \mathcal{L}\varepsilon, \\
&\mathcal{L} = \frac{1}{2}\omega \left\{ \omega \langle \boldsymbol{A}, \bar{\boldsymbol{A}} \rangle \langle \dot{\mathbf{x}}, \dot{\mathbf{x}} \rangle + \left(\mathrm{i} \left\langle \boldsymbol{A}, \dot{\bar{\boldsymbol{A}}} \right\rangle + \text{C.C.} \right) \right\}.
\end{aligned} \tag{10}$$

6. Helicity of Constrained Fields

The action integral is evidently invariant under rotations of the frame in the plane formed by the vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$. This invariance yields some conservation law due to the Noether theorem. To find the law we consider the change of the form of the action integral under rotations of the local frames specified by an infinitesimal matrix $\delta\eta^a_b(s)$ defined as a function of the parameter s on each curve. Rotation of the frame changes components of the vectors $\boldsymbol{A}$, $\bar{\boldsymbol{A}}$ but does not change the vectors. Thus, the first term of the Lagrangian (10) containing scalar product $< \boldsymbol{A}, \bar{\boldsymbol{A}} >$ does not contribute to the variation. At the same time the rotation changes the derivative of $\dot{\boldsymbol{A}}$:

$$\begin{aligned}
\delta \dot{\boldsymbol{A}}^a &= \delta \left(\frac{\mathrm{d}A^a}{\mathrm{d}s} + \gamma_{bc\cdot}{}^a \dot{x}^b A_c \right) \\
&= \left(\frac{D\delta A^a}{\mathrm{d}s} \right) + \delta\omega_{c\cdot}{}^a(\dot{\mathbf{x}}) A^c \boldsymbol{e}_a, \\
\delta\omega_{c\cdot}{}^a &= D_b(\delta\eta_{c\cdot}{}^a)\, \theta^b,
\end{aligned}$$

where the covariant derivative $D_b(\delta\eta^a_{c\cdot})$ is exactly the variation of the connected 1-form.

The first term in the variation of $\dot{\boldsymbol{A}}$ does not contribute to the variation of the action due to the field equations. Thus the variation of the action is:

$$\begin{aligned}
\int \delta\mathcal{L}\varepsilon &= \int \left\{ \mathrm{i}\omega \left\langle \boldsymbol{A}, \frac{D\delta\eta_{a\cdot}{}^b}{\mathrm{d}s} \bar{A}^a \boldsymbol{e}_b \right\rangle + \text{C. C.} \right\} \varepsilon \\
&= \int \left\{ \mathrm{i}\omega\, (\delta\eta_{ab})^\bullet \bar{A}^a A^b + \text{C. C.} \right\} \varepsilon \\
&= \int \mathrm{d}[\ldots] + \int 2\omega \delta\eta_{ab} \frac{D}{\mathrm{d}s} \left(\frac{\bar{A}^a A^b - \bar{A}^b A^a}{2\mathrm{i}} \right) \varepsilon.
\end{aligned}$$

Finally, we have:

$$\frac{DS_{ab}}{\mathrm{d}s} = 0.$$

So, due to the Noether theorem we obtain the conserved spin current with a single non-zero component

$$J^-_{ab} = J_{+ab} = 2S_{ab}, \tag{11}$$

where

$$\boldsymbol{S}^{ab} = \frac{\omega}{2\mathrm{i}} \left(\bar{A}^a A^b - \bar{A}^b A^a \right)$$

is the spin tensor of the wave. Due to the constraint (6) it is zero for linearly polarized waves, while for circularly polarized waves it has a single non-zero element

$$\boldsymbol{S}^{12} = \pm\omega |a^1 a^2|,$$

whose sign depends only on helicity. The only consequence of this result which we need is, that the spin has one single non-zero component. As for its magnitude, we accept its quantum value 1 in dimensionless units. Note that the spin always points along the vector of velocity, so no special equation is needed for it. This fact provides the implementation of the Tulczyjew constraint which requires that the conversion of the particle momentum with its spin is zero [5].

7. Field Equations

Now we return to the action functional (10) and compute its variations under small variations of the field which has two components A_1 and A_2. The variation of the first term in the constrained Lagrangian is identacally zero because it contains the factor $\langle\theta^+,\theta^+\rangle$ which is not varied, therefore we ignore it. The variation of the rest part of the action is:

$$\begin{aligned}\delta\mathcal{A} &= \delta\int\left\{\left\langle \boldsymbol{e}_+\circ\boldsymbol{A}, \boldsymbol{e}_-\circ\bar{\boldsymbol{A}}\right\rangle + \mathrm{C.\,C.}\right\}\varepsilon \\ &= \int\Big[\left\langle\delta\left(\boldsymbol{e}_+\circ\boldsymbol{A}\right), \boldsymbol{e}_-\circ\bar{\boldsymbol{A}}\right\rangle \\ &\qquad + \left\langle \boldsymbol{e}_+\circ\boldsymbol{A}, \delta\left(\boldsymbol{e}_-\circ\bar{\boldsymbol{A}}\right)\right\rangle + \mathrm{C.\,C.}\Big]\varepsilon \\ &= \int\Big[\left\langle \boldsymbol{e}_+\circ\delta\boldsymbol{A}, \boldsymbol{e}_-\circ\bar{\boldsymbol{A}}\right\rangle \\ &\qquad + \left\langle \boldsymbol{e}_+\circ\boldsymbol{A}, \boldsymbol{e}_-\circ\delta\bar{\boldsymbol{A}}\right\rangle + \mathrm{C.\,C.}\Big]\varepsilon,\end{aligned}$$

where the action of the vectors $\boldsymbol{e}_\pm$ is considered as differentiation along the vectors. The next step is to extract the total derivatives:

$$\begin{aligned}\delta\mathcal{A} = \int &\left[\boldsymbol{e}_+\circ\left\langle\delta\boldsymbol{A}, \boldsymbol{e}_-\circ\bar{\boldsymbol{A}}\right\rangle\right. \\ &\quad \left. + \boldsymbol{e}_-\circ\left\langle \boldsymbol{e}_+\circ\boldsymbol{A}, \delta\bar{\boldsymbol{A}}\right\rangle + \mathrm{C.\,C.}\right]\varepsilon \\ &- \int\left\langle\delta\bar{\boldsymbol{A}},\ \boldsymbol{e}_+\circ\left(\boldsymbol{e}_-\circ\boldsymbol{A}\right) + \boldsymbol{e}_-\circ\left(\boldsymbol{e}_+\circ\boldsymbol{A}\right)\right\rangle\varepsilon \\ &- \int\left\langle\delta\boldsymbol{A},\ \boldsymbol{e}_+\circ\left(\boldsymbol{e}_-\circ\bar{\boldsymbol{A}}\right) + \boldsymbol{e}_-\circ\left(\boldsymbol{e}_+\circ\bar{\mathbf{A}}\right)\right\rangle\varepsilon.\end{aligned}$$

Note that due to the metric of the null-frame (2) combinations like $\boldsymbol{e}_+\circ V_-$ are parts of the divergence of a vector $\mathbf{V}$, and if the vector has only one component '$-$', this expression coincides with its divergence. In particular, the combination $\boldsymbol{e}_{\{+}\circ\left(\boldsymbol{e}_{-\}}\circ f\right)$ is exactly the D'Alembert operator applied to a function f, which is constant on the wave fronts. Consequently, the first term in the right-hand side of the equation above is exactly a divergence, hence the integral can be taken by parts and the variation of the action integral becomes:

$$\int \mathrm{d}[\ldots] - \int\left\{\left\langle\delta\bar{\boldsymbol{A}},\ \boldsymbol{e}_{\{+}\circ\left(\boldsymbol{e}_{-\}}\circ\boldsymbol{A}\right)\right\rangle + \mathrm{C.\,C.}\right\}\varepsilon.$$

Here the first term reduces to a surface integral and the brackets at the subscripts mean symmetrization. As usual, the surface integral vanishes at infinity and all the rest reduces to the following Euler-Lagrange equation:

$$\boldsymbol{e}_{\{+}\circ\left(\boldsymbol{e}_{-\}}\circ\boldsymbol{A}\right) = 0.$$

Evidently, this covariant equation reduces to the D'Alembert equation in local Cartesian coordinates, provided that the curves $x(s)$ are locally null straight lines. In fact, any vector whose covariant derivative along the curve is zero:

$$\boldsymbol{e}_-\circ\boldsymbol{A} = 0, \tag{12}$$

and so for the complex conjugate, satisfies this equation. The constrained fields satisfy this equation due to equations (1 – 6), consequently, this part of the entire variation of the action integral is zero due to the constraints. Though the field equation leaves ω to be an arbitrary function of the phase we restrict our analysis with monochromatic waves, for which this value is constant.

8. Papapetrou Equations

It remains to consider the second part of the variation of the action integral, produced by the variation of the local frames under fixed field variables. Since the local frames are defined as co-moving frames on the congruence of null-curves, it is possible to introduce the variation of the congruence and to derive the variation of the frames from this. A small change of the shape of a curve $x(s)$ causes a small change of the tangent vector direction without a change of its length. Consequently, the variation of the vector $\boldsymbol{e}_-$ is orthogonal to it and to the complementary null-vector $\boldsymbol{e}_+$, hence, belongs to the span of the two polarization vectors $\boldsymbol{e}_\beta$. Since the vector $\boldsymbol{e}_+$ is also orthogonal to the span, it suffers no change. Therefore, the variations of both $\boldsymbol{e}_+$ and of the corresponding covector θ^- are zero. Thus, only variations of the vector $\boldsymbol{e}_-$ and the corresponding 1-form θ^+ contribute to the variation of the action integral (10).

The first term in the Lagrangian (10) contains the factor $\langle\theta^+,\theta^+\rangle = \langle \boldsymbol{e}_-,\boldsymbol{e}_-\rangle \equiv 0$, therefore its contribution is predetermined only by the variation of the vector $\boldsymbol{e}_- = \dot{\mathbf{x}}$. Consider the variation of the four-dimensional integral

$$\frac{\omega^2}{2}\int |\boldsymbol{A}|^2 \langle \dot{\boldsymbol{x}},\dot{\boldsymbol{x}}\rangle\,\varepsilon$$

with respect to the variation of the curves. The variation of the factor $\langle \dot{\boldsymbol{x}},\dot{\boldsymbol{x}}\rangle$ is well-known from the variation principle for geodesics [6]. Here we can use the fact that the variation of the integrand reduces to the scalar product of the vector of variation of the curve $\delta\boldsymbol{x}(s)$ and the covariant acceleration $\frac{D\dot{\boldsymbol{x}}}{ds}$:

$$\delta\int \frac{\omega^2}{2}|\boldsymbol{A}|^2 \langle \dot{\boldsymbol{x}},\dot{\boldsymbol{x}}\rangle\,\varepsilon = \omega^2\int |\boldsymbol{A}|^2 \left\langle \delta\boldsymbol{x}(s), \frac{D\dot{\boldsymbol{x}}}{\mathrm{d}s}\right\rangle. \quad (13)$$

The second term in the Lagrangian (10) has a single zero factor $\breve{\boldsymbol{A}}$ in the scalar product, consequently, a non-zero contribution to the variation of the action integral appears only when varying this factor. The zero factor to be varied is $\breve{\boldsymbol{A}}$:

$$\breve{\boldsymbol{A}} = \left(\frac{\mathrm{d}\bar{A}^a}{\mathrm{d}s} + \gamma_{bc\cdot}{}^{a}\dot{x}^b\bar{A}^c\right)\boldsymbol{e}_a.$$

The variation of the covariant derivative contains the derivative with respect to s and the term containing the connection $\gamma_{bc\cdot}{}^{a}$. The variation of the first of them does not contrtibute to the variation of the action integral because neither the polarization vectors $\boldsymbol{e}_\alpha$ nor the components of the field change under varying the congruence of curves. The only term which is affected from the change is the connection $\gamma_{bc\cdot}{}^{a}$. Therefore we can write the contribution of this term as follows:

$$\frac{1}{2}\delta\int \omega\left(\mathrm{i}\left\langle \boldsymbol{A},\breve{\boldsymbol{A}}\right\rangle + \text{C. C.}\right)\varepsilon = -\frac{1}{2}\int\left(\mathrm{i}A^b\delta\,\gamma_{bc\cdot}{}^{a}\bar{A}^c + \text{C. C.}\right)\varepsilon.$$

Here the minus sign appears because of the lower index at the field component. Thus, the next task is to find the variation of the connection $\delta\,\gamma_{bc\cdot}{}^{a}$.

It is easier to find the variation of the 1-form $\omega_{a\cdot}{}^{b}$ because this has to be taken in a fixed point under changing the field of frames which is dragged by the vector $\delta\boldsymbol{x}$. This variation is, by definition, the Lie derivative of the connection 1-form with respect to this vector. So, to find the variation of the connection 1-form it suffices to take its Lie derivative with respect to the vector $\delta\boldsymbol{x}$. The Lie derivative of a 1-form λ with respect to a vector $\vec{v}$ is [7]:

$$\pounds_{\vec{v}}\lambda = \mathrm{d}\lambda(\vec{v}) + \mathrm{d}(\lambda(\vec{v})).$$

Unlike the ordinary 1-form the connection has components referring to local frames. Since the frames are built on the vector of velocity on the congruence, its variation causes some infinitesimal rotation of the frames. Denote the corresponding matrix of rotation $\eta_a{}^c$. This rotation transforms components of the connection 1-form and must be taken into account. To do this it is necessary to calculate the explicit form of this matrix from the dragging vector $\delta\boldsymbol{x}$.

Consider a point in space-time, a curve passing through it and the local frame and the small variation of the congruence of curves given by the small vector $\delta\boldsymbol{x}$ which drags the congruence. This dragging replaces the curve passing through this point with the curve dragged by the vector, and the local frame built in this point is also to be replaced by the frame dragged by this vector from a neighboring point. Since, on the one hand, both the frames are orthonormal, the variation of the frames is a small rotation. On the other hand, since this rotation is specified by dragging an orthonormal frame from a neighboring point, this transformation is given by the connection itself. In other words, the matrix of rotation $\eta_a{}^c$ is exactly the value of the form of connection on the dragging vector: $\eta_a{}^b = \omega_{a\cdot}{}^{b}(\delta\boldsymbol{x})$.

Thus, the Lie derivative of the connection 1-form with respect to the vector $\delta\boldsymbol{x}$ is:

$$\pounds_{\delta\boldsymbol{x}}\omega_{a\cdot}{}^{b} = d\omega_{a\cdot}{}^{b}(\delta\boldsymbol{x}) + d(\omega_{a\cdot}{}^{b}(\delta\boldsymbol{x})) - \eta_a{}^c\omega_{c\cdot}{}^{b} + \eta_c{}^b\omega_{a\cdot}{}^{c}.$$

Subsituting the matrix of rotation we obtain the desired Lie derivative:

$$\begin{aligned}\pounds_{\delta\boldsymbol{x}}\omega_{a\cdot}{}^{b} &= d\omega_{a\cdot}{}^{b}(\delta\boldsymbol{x}) - \omega_{a\cdot}{}^{c}(\delta\boldsymbol{x})\omega_{c\cdot}{}^{b} \\ &\quad + \omega_{c\cdot}{}^{b}(\delta\boldsymbol{x})\omega_{a\cdot}{}^{c} + d(\omega_{a\cdot}{}^{b}(\delta\boldsymbol{x})) \\ &= (d\omega_{a\cdot}{}^{b}(\delta\boldsymbol{x}) + \omega_{c\cdot}{}^{b}\wedge\omega_{a\cdot}{}^{c})(\delta\boldsymbol{x}) + d(\omega_{a\cdot}{}^{b}(\delta\boldsymbol{x})) \\ &= \Omega_a{}^b(\delta\boldsymbol{x}) + d(\omega_{a\cdot}{}^{b}(\delta\boldsymbol{x})).\end{aligned}$$

Thus we have obtaind the curvature 2-form $\Omega_a{}^b \equiv R_{cd\underline{a}}{}^b\theta^c\wedge\theta^d$. Substituting now this into the variation of $\boldsymbol{A}$ gives:

$$\begin{aligned}\delta\breve{\boldsymbol{A}} &= R_{dbc}{}^{a}\delta x^d\dot{x}^b\bar{A}^c\boldsymbol{e}_a + \dot{x}^d\partial_d(\gamma_{bc\cdot}{}^{a}\,\delta x^b\bar{A}^c)\boldsymbol{e}_a \\ &= R_{dbc}{}^{a}\delta x^d\dot{x}^b\bar{A}^c\boldsymbol{e}_a + (\gamma_{bc\cdot}{}^{a}\,\delta x^b\bar{A}^c)^{\bullet}\boldsymbol{e}_a.\end{aligned}$$

The variation of $\dot{\boldsymbol{A}}$ is similar, so after composing the total variation of the second term in the action integral we obtain two terms. One is the total derivative of $\omega(A_a \gamma_{bc\cdot}{}^a \delta x^b \bar{A}^c)$ with respect to s, which vanishes on the endpoints of the curves. Thus, the total variation of this part of the Lagrangian is given by another term which is:

$$\frac{\mathrm{i}\omega}{2} R_{dbc}{}^a \delta x^d \dot{x}^b (\bar{A}^c A_a - \bar{A}_a A^c) = -R_{dbc}{}^a \delta x^d \dot{x}^b S^c{}_a,$$

where we have introduced the spin by its only component $S^1{}_2$. The Euler-Lagrange equation for the curves $x(s)$ coincides with the Papapetrou equation:

$$\omega^2 |\boldsymbol{A}|^2 \frac{D\dot{x}^a}{\mathrm{d}s} = R^a_{db\cdot c} \dot{x}^c S^{db}. \tag{14}$$

9. Conclusion

A geometric-optical description of an electromagnetic wave in curved space-time, which account for its polarization is obtained straightforwardly from the classical variational principle for the electromagnetic field. For this purpose the full functional space of the electromagnetic field is reduced to its subspace of locally plane monochromatic waves. Therefore, first of all, the notion of locally plane monochromatic waves in curved space-time must be defined. It turns out that waves of this sort exist, provided that their wavelengths are small compared with the scales under consideration. Assuming this, we have formulated the constraints under which the full functional space of the electromagnetic field reduces to its subspace of locally plane monochromatic waves and impose this as constraints. These constraints not only reduce the field variables but also introduce variables of another kind which specify a field of local frames associated to the wave, containing some congruence of null-curves $x^i(s)$. These curves become the main object in the construction because they specify the field of local frames, and the field variables $\dot{\boldsymbol{A}}$ and $\bar{\boldsymbol{A}}$ are referring to this frame.

Returning to the action principle for the constrained electromagnetic field we have obtained the Lagrangian (10), which contains variables of two kinds, namely a congruence of curves $x^i(s)$ and the field itself. Thus we have two kinds of Euler-Lagrange equations. The equations of first kind reduce to the local D'Alembert equation for the field components $\dot{\boldsymbol{A}}$ and $\bar{\boldsymbol{A}}$ which are trivial due to the constraints imposed. The variation of the curves yields all the other equations which contain the main result of this investigation. It turns out that the corresponding Euler-Lagrange equations are exactly the Papapetrou equations for a classical massless particle with helicity 1. This equation determines the shape of the 0-curves which, by construction, can be considered as world lines of photons with the same wavelengths and helicities. They apparently differ from null-geodesics and, thereby manifest the influence of spin-gravitational interaction on the propagation of the electromagnetic waves in gravitational fields. The effect of this interaction is proportional to the wavelength [4], therefore this fact can be observed in radioastronomy.

[1] B. Mashhoon, J. Math. Phys. **12**, 1075 (1971).

[2] A. Frydryszak, Lagrangian Models of Particles with Spin: the First Seventy Years. arXive:hep-th/9601020.

[3] Z. Ya. Turakulov and M. Safonova, Mod. Phys. Lett. A **18**, 579 (2003).

[4] Z. Ya. Turakulov and M. Safonova, Mod. Phys. Lett. A **20**, 2785 (2005).

[5] W. Tulczyjew, Acta Phys. Pol. **18**, 393 (1959).

[6] C. W. Misner, K. S. Thorn, and J. A. Wheeler, Gravitation, Freeman, San Francisco 1973.

[7] B. F. Schutz, Geometrical Methods of Mathematical Physics, Cambridge University Press, Cambridge 1980.