International Journal of Theoretical Physics, Vol. 43, No. 10, October 2004 (© 2004)

Derivation of the Bialynicki-Birula Photon Wave Function From Three-Component Spinors

Jaroslaw Zaleśny[1]

A simple derivation of the Bialynicki-Birula photon wave function equation (equivalent to the Maxwell equations) in the formalism of the three-component spinors has been presented. The derivation has been based on two assumptions: (1) The relativistic energy–momentum relation for a massless particle is satisfied; (2) The description of a spin in three-dimensional spinor spaces S and $\dot{S}$ is the same in the limit when velocity of the inertial frame moving along the photon propagation direction approaches c.

KEY WORDS: photon wave function; Maxwell equations; spinors.

1. INTRODUCTION

As it is well known the derivation of the Dirac equation may be based on two independent statements (e.g., Brzezowski, 1995; Lopuszański, 1985):

- the relativistic energy–momentum relation for a massive particle must be satisfied;
- the description of a spin in two-dimensional spinor spaces S and $\dot{S}$ (dotted S) should be the same in the inertial frame in which the particle is in rest.

The Dirac equation is a compact record of the information contained in these statements. The group $SL(2, C)$ represented by 2×2 complex matrices with unit determinant acts in S and $\dot{S}$. The subspace of Hermitian tensors of $S \otimes \dot{S}$ can be identified with the space of relativistic four-vectors. In this way a close relation between the group $SL(2, C)$ and the Lorentz group is achieved.

In this paper, I intend to show that in the similar way it is possible to construct the Bialynicki-Birula photon wave function equation (Bialynicki-Birula, 1994). I will follow the analogy with the derivation of the Dirac equation as much as possible. I will underline the similarities and differences.

There is a vast literature devoted to the spinor formulation of Dirac and Maxwell equations. I am not able to give here a comprehensive account. Let

[1] Institute of Physics, Technical University of Szczecin, Al. Piastów 48, 70-310, Szczecin, Poland; e-mail: jarek@ps.pl.

me restrict myself to mention only a few references. Among the most early are perhaps (Laporte, 1931; Oppenheimer, 1931; Good, 1957; Moses, 1959). The reader may find a comprehensive account on the Bialynicki-Birula photon wave function in (Bialynicki-Birula, 1996) and (Kobe, 1999). An interesting discussion on the "equivalence" of the Maxwell and Dirac equation could be found (e.g., in Vaz and Rodrigues, 1995; Gsponer, 2002).

2. DERIVATION

In contrast to the case of Dirac equation, where the main object of interest are two-dimensional complex vectors (i.e., Pauli spinors or simply spinors), we are now interested in three-dimensional complex vectors (three component spinors). For simplicity, from now on, I will call them also spinors. The spinors form a vector space S.

$$|\xi\rangle \in S, \quad |\xi\rangle = \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix} = \begin{pmatrix} E_x + iH_x \\ E_y + iH_y \\ E_z + iH_z \end{pmatrix}. \tag{1}$$

The real quantities E and H have no physical interpretation as yet. They are simply real and imaginary parts of the spinor. Apart from the space S we need also its duplicate: the space $\dot{S}$ (dotted S).

In the case of Dirac equation the six-parameter $SL(2, C)$ group acts in two-dimensional spinor spaces. Therefore, it is possible to find a relation of the $SL(2, C)$ to the six-parameter Lorentz group. In the present case, we must now decide what a group acts in the spaces of three-component spinors. We also need a relation of the group to the Lorentz group, therefore it has to be a six-parameter group. This requirement fulfils the $O(3, C)$ group represented by complex 3×3 matrices U

$$U^{\mathrm{T}}U = 1, \tag{2}$$

the superscript T means a transpose. The matrices U act in S and $\dot{S}$ in the following way

$$U|\xi\rangle = |\xi'\rangle, \quad \bar{U}|\dot{\xi}\rangle = |\dot{\xi}'\rangle, \tag{3}$$

the bar over U denotes a complex conjugation. Note that

$$|\xi'\rangle^{\mathrm{T}}|\xi'\rangle = (U|\xi\rangle)^{\mathrm{T}}U|\xi\rangle = |\xi\rangle^{\mathrm{T}}U^{\mathrm{T}}U|\xi\rangle = |\xi\rangle^{\mathrm{T}}|\xi\rangle = E^2 - H^2 + 2i\mathbf{E}\cdot\mathbf{H} \tag{4}$$

is an invariant of the transformation, whereas

$$\langle\xi|\xi\rangle = E^2 + H^2, \tag{5}$$

where $\langle\xi| \equiv |\bar{\xi}\rangle^{\mathrm{T}}$. Let's consider a tensor product $S \otimes \dot{S}$. We are interested only in the subset of Hermitian tensors because the $O(3, C)$ group acts inside the subset, i.e. after the transformation the Hermitian tensor T remains Hermitian

$$T' = UT\bar{U}^{\mathrm{T}}. \tag{6}$$

In the case of Dirac equation the 2×2 Hermitian tensors can be identified with Lorentz four-vectors. In the present case the Hermitian condition is too weak. The 3×3 Hermitian tensors have too many parameters to be identified with relativistic four-vectors. Therefore, we impose on the tensors one more condition reducing the number of parameters to three

$$\frac{1}{2}(T + \bar{T}) = 1. \tag{7}$$

The Hermitian tensors satisfying the condition have the following explicit form

$$T \equiv T_{ij} = \begin{pmatrix} 1, & ix^3, & -ix^2 \\ -ix^3, & 1, & ix^1 \\ ix^2 & -ix^1, & 1 \end{pmatrix}, \tag{8}$$

and the determinant

$$\det T = 1 - (x^1)^2 - (x^2)^2 - (x^3)^2. \tag{9}$$

Certainly, the tensors cannot be identified with four-vectors. However, one may identify them with three-dimensional unit vectors. To this end, I impose on the determinant the condition

$$\det T = 0. \tag{10}$$

Therefore the quantities x^1, x^2, x^3 may be interpreted as direction cosines. Later I will interpret the unit vector as a vector aiming at the direction of the propagating photon in an inertial frame. The condition (7) and the tensors (8) are not Lorentz-invariant, i.e. they are not invariant under the transformation (6). However, they do not change their form under the following transformation

$$T' = RTR^{\mathrm{T}}, \quad \text{where} \quad R^{\mathrm{T}}R = 1. \tag{11}$$

The real three-parameter matrices R form the subgroup $O(3, R)$ of the $O(3, C)$ matrices U. Note that the transformations (11) as well as the transformations (6) leave the determinant of T unchanged. Note also that the tensors (8) can be written in the following form

$$T = 1 - x^1 S_1 - x^2 S_2 - x^3 S_3, \tag{12}$$

where 1 means here a 3×3 unit matrix and S_k, $k = 1, 2, 3$ are generators of the rotation group, $(S_k)_{ij} = i\varepsilon_{kij}$, where ε_{kij} is the antisymmetric Levi-Civita symbol.

The Dirac equation may be derived under the assumption that the pictures of a spin in two-dimensional spaces S and $\dot{S}$ are identical in the inertial frame in which a massive particle is in rest. Here, we try to describe photon. And for the photon there is no such inertial frame in which it rests. The transformations (3) transform spinors from one inertial frame to another but no frame exists in which the pictures of the spin in S and $\dot{S}$ are the same. Otherwise, the frame would be distinguished, in contrast to the relativistic theory. However, we may choose the following solution. Imagine that in an inertial frame we observe photon propagating with the velocity c. We may describe the direction of the propagation with the help of the direction cosines x^1, x^2, x^3. Now let us try to make a transformation of the spinor ξ to "the frame moving with velocity c" in the direction pointed by the unit vector (x^1, x^2, x^3). The transformation has to be understood in a sense of a limit. Therefore, the transformation cannot be made by the $O(3, C)$ matrices. It has to be a 3×3 matrix depending on the three parameters x^1, x^2, x^3. Moreover, if the frame in which we observe the photon is rotated, the unit vector would change to $(x^{1\prime}, x^{2\prime}, x^{3\prime})$ but the form of the matrix has to be the same because of isotropy of space. All these requirements fulfils the matrix T. What could be the result of the very special transformation of the spinor $|\xi\rangle$? According to the special relativity theory, in the inertial frame moving with the velocity c, time stops. If the spin means something like "spinning," then in this special frame the spin has to vanish (no time, no motion). Therefore, as an equivalence of the condition of "the same spin pictures in S and $\dot{S}$" I propose for photon the following conditions

$$T|\xi\rangle = 0, \quad \bar{T}|\dot{\xi}\rangle = 0. \tag{13}$$

Note that the condition (10) is in agreement with the requirement that the equations must have nontrivial solutions. Taking complex conjugation of the second equations of (13) and comparing it with the first one of the pair one may conclude that

$$|\dot{\xi}\rangle = |\bar{\xi}\rangle. \tag{14}$$

The two equations of (13) can be written as a single one

$$\begin{pmatrix} T & 0 \\ 0 & \bar{T} \end{pmatrix} \left| \begin{matrix} \xi \\ \hline \bar{\xi} \end{matrix} \right\rangle = 0. \tag{15}$$

We have assumed that the spin state of the photon is described by the bispinor

$$\left| \begin{matrix} \xi \\ \hline \bar{\xi} \end{matrix} \right\rangle. \tag{16}$$

As a matter of fact it is sufficient to solve one of the equations of (13), e.g., the first one in order to find $|\xi\rangle$, and then to find $|\dot{\xi}\rangle$ from (14). Note, however, that the whole construction requires in fact the "doubled" equation (15).

The equation $T|\xi\rangle = 0$ may be written in the form

$$(1 - x^1 S_1 - x^2 S_2 - x^3 S_3)|\xi\rangle = 0. \tag{17}$$

We remember that the energy–momentum relation for zero mass particle is the second foundation stone of our construction. Therefore, we can multiply (17) by the magnitude of a three-dimensional momentum vector of the photon $p = \mathcal{E}/c$. Because the quantities x^1, x^2, x^3 are direction cosines, therefore $px^1 = p^1$, $px^2 = p^2$, $px^3 = p^3$ are components of the three-dimensional momentum $\mathbf{p}$. Note also that the condition (10)

$$(x^1)^2 + (x^2)^2 + (x^3)^2 = 1, \tag{18}$$

after multiplication by $p^2 = \mathcal{E}^2/c^2$ reproduces the energy–momentum relation for a zero mass particle

$$(p^1)^2 + (p^2)^2 + (p^3)^2 = \frac{\mathcal{E}^2}{c^2}. \tag{19}$$

Thus, we may write the Equation (17) in the form

$$\frac{\mathcal{E}}{c}|\xi\rangle = \mathbf{p} \cdot \mathbf{S}|\xi\rangle. \tag{20}$$

The equation can be rewritten in a vector notation as

$$\frac{\mathcal{E}}{c}|\xi\rangle = i\mathbf{p} \times |\xi\rangle \tag{21}$$

using the mathematical identity

$$(\mathbf{a} \cdot \mathbf{S})\mathbf{b} = i\mathbf{a} \times \mathbf{b}. \tag{22}$$

($\mathbf{a}$, $\mathbf{b}$ are any three-dimensional vectors). Iterating (20) or (21) and taking into account that the energy–momentum relation $\mathcal{E}^2/c^2 = p^2$ is satisfied, an additional condition on the spinor $|\xi\rangle$ is obtained

$$\mathbf{p} \cdot |\xi\rangle = 0. \tag{23}$$

Equations (21) and (23) are in fact vacuum Maxwell equations. We may call them "classical" Maxwell equations, because they describe classical relativistic massless particle with some "internal" structure (spin). Up to now there is nothing "wave" or "quantum" in them. To obtain the well known usual Maxwell equations we use Equation (1) and interpret the real and imaginary part of the spinor as an electric and magnetic field, respectively

$$|\xi\rangle = \mathbf{E} + i\mathbf{H}. \tag{24}$$

Applying the "recipe for quantum mechanics"

$$\mathcal{E} \rightarrow i\hbar\frac{\partial}{\partial t}, \quad \mathbf{p} \rightarrow -i\hbar\nabla, \tag{25}$$

and taking real and imaginary parts of (21) and (23) one can obtain from (21)

$$\partial_t\mathbf{E} = c\nabla \times \mathbf{H}, \quad \partial_t\mathbf{H} = -c\nabla \times \mathbf{E}, \tag{26}$$

and from (23)

$$\nabla \cdot \mathbf{E} = 0, \quad \nabla \cdot \mathbf{H} = 0. \tag{27}$$

From the above procedure it is quite clear that the usual Maxwell equations are in fact quantum equations. It is simply a joke of Nature that the constant $\hbar$ cancels out. It is a consequence of the zero mass of photon. And because photons are bosons, the interference picture may be macroscopically observed.

Each of the Equations (13) separately leads to the Maxwell equations (26) and (27). However, the full description of the photon spin requires the Equation (15). After multiplying it by $p = \mathcal{E}/c$ and after some obvious manipulations one obtains the Bialynicki-Birula photon wave function equation

$$i\hbar\, \partial_t \psi = c \begin{pmatrix} \mathbf{p} \cdot \mathbf{S} & 0 \\ 0 & -\mathbf{p} \cdot \mathbf{S} \end{pmatrix} \psi, \tag{28}$$

where the photon wave function

$$\psi = \binom{\mathbf{E}+i\mathbf{H}}{\mathbf{E}-i\mathbf{H}} \quad \text{fulfils additional condition} \quad \mathbf{p} \cdot \psi = 0, \tag{29}$$

here $\mathbf{p} \equiv -i\hbar\nabla$.

This equation may be regarded as a Schrödinger equation for photon.

3. ANOTHER POINT OF VIEW

The most important points in the derivation are Equations (8) and (13). If we seek the equation for photon which is linear in operators of energy and momentum, then the form (8) of tensor $T = 1 - \hat{\mathbf{x}} \cdot \mathbf{S}$ is unique with an accuracy of some trivial transformations. Note, however, that because of the condition (29) (or (23), which states that $\hat{\mathbf{x}} \cdot |\xi\rangle = 0$, one may complicate the form of T without any consequence, e.g.,

$$T' = 1 - \hat{\mathbf{x}} \cdot \mathbf{S} + (\hat{\mathbf{x}} \cdot \mathbf{S})(\hat{\mathbf{x}} \cdot \mathbf{S}) - \hat{\mathbf{x}} \cdot \hat{\mathbf{x}}. \tag{30}$$

Tensors T' and T are indiscernible in action on spinor, because of the following mathematical identity

$$[\hat{\mathbf{x}} \cdot \hat{\mathbf{x}} - (\hat{\mathbf{x}} \cdot \mathbf{S})(\hat{\mathbf{x}} \cdot \mathbf{S})]|\xi\rangle = \hat{\mathbf{x}}(\hat{\mathbf{x}} \cdot |\xi\rangle). \tag{31}$$

Moreover, $\det T' = (1 - (x^1)^2 - (x^2)^2 - (x^3)^2)^2$. Therefore, one may use T' instead of T from the beginning. As earlier, the condition (23) follows then from iteration as a result of energy–momentum relation. And this allows to cut down T' to T.

Let us present now apparently another approach to the derivation of vacuum Maxwell equations. Once more we base our considerations on the relativity theory. Suppose that all we know is that the electromagnetic field is described by an antisymmetric tensor $F^{\mu\nu}$. The electric **E** and magnetic **H** fields are simply components of the tensor. They are transformed from one inertial frame to another in the following way (Jackson, 1975)

$$\mathbf{E}' = \gamma(\mathbf{E} + \vec{\beta} \times \mathbf{H}) - \frac{\gamma^2}{\gamma + 1}\vec{\beta}(\vec{\beta} \cdot \mathbf{E})$$

$$\mathbf{H}' = \gamma(\mathbf{H} - \vec{\beta} \times \mathbf{E}) - \frac{\gamma^2}{\gamma + 1}\vec{\beta}(\vec{\beta} \cdot \mathbf{H}). \tag{32}$$

The fields **E** and **H** are not independent, therefore it is useful to introduce the Silberstein complex vector (Silberstein, 1907)

$$\mathbf{F} = \mathbf{E} + i\mathbf{H}. \tag{33}$$

It contains entire information about the electromagnetic field. As a matter of fact the Silberstein vector is nothing else as the spinor $|\xi\rangle$ (24). It transforms in the following way

$$\mathbf{F}' = \gamma\mathbf{F} - i\gamma\vec{\beta} \times \mathbf{F} - \frac{\gamma^2}{\gamma + 1}\vec{\beta}(\vec{\beta} \cdot \mathbf{F}). \tag{34}$$

With the help of the identity (22) it can be written as

$$\frac{1}{\gamma}\mathbf{F}' = \left[(1 - \vec{\beta} \cdot \mathbf{S}) + \frac{\gamma}{\gamma + 1}((\vec{\beta} \cdot \mathbf{S})(\vec{\beta} \cdot \mathbf{S}) - \beta^2)\right]\mathbf{F}. \tag{35}$$

Do not forget, however, that we want to describe photon—the particle moving with velocity c. Let the direction of the photon in some inertial frame is given by the direction cosines x^1, x^2, x^3 and let the new inertial frame moves in the direction with velocity β. Now go to the limit $\beta \to 1$. In that limit $\beta^1 \to x^1$, $\beta^2 \to x^2$, $\beta^3 \to x^3$, $1/\gamma \to 0$ and $\gamma/(\gamma + 1) \to 1$. We suppose also that in this limit F$'$ vanishes. Thus, from (35) one obtains

$$T'\mathbf{F} = 0, \tag{36}$$

where T' is given by (30). As has been already discussed it may be cut down to

$$T\mathbf{F} = 0. \tag{37}$$

This is equivalent to the first of equations of (13). The second one could be obtained starting with complex conjugation of the Silberstein vector. The further procedure allowing to obtain the Maxwell equations is the same as already described in the previous section.

4. SUMMARY

In this paper the simple derivation of the Bialynicki-Birula wave function equation (equivalent in some sense to the vacuum Maxwell equations) has been presented in the language of the three-component spinors. It has been derived on the basis of the relativistic energy–momentum relation for a massless particle and the Equation (13). The heuristic arguments leading to the Equation (13) someone may find as controversial, however, the efficiency of this equation has been demonstrated in the paper. Up to my knowledge the presented derivation is original.

REFERENCES

Bialynicki-Birula, I. (1994). *Acta Physics Polonica Part A* **86**, 97–116; http://www.cft.edu.pl/ birula/publ.html

Bialynicki-Birula, I. (1996). In *Progress in Optics*, *Vol. 36*, E. Wolf, ed., Elsevier, Amsterdam, pp. 245–294; http://www.cft.edu.pl/ birula/publ.html

Brzezowski, S. (1995). *Spinory (Spinors, in Polish)*, Uniwersytet Jagielloński, Kraków.

Good, R. H. (1957). *Physics Review* **105**, 1914–1919.

Gsponer, A. (2002). *International Journal of Theoritical Physics* **41**, 689–694; http://www.arXiv.org/abs/math-ph/0201053

Jackson, J. D. (1975). *Classical Electrodynamics*, Wiley, New York.

Kobe, D. H. (1999). *Foundations of Physics* **29**(8), 1203–1231.

Laporte, O. and Uhlenbeck, G. (1931). *Phys. Rev.* **37**, 1380–1397.

Lopuszański, J. (1985). *Rachunek spinorów (Spinor calculus, in Polish)*, PWN, Warszawa.

Moses, H. E. (1959). *Physical Review* **113**(6), 1670–1679.

Oppenheimer, J. R. (1931). *Physics Revew* **38**, 725–746.

Silberstein, L. (1907). *Annals der Physics* **22**(579); *Annals der Physics* **24**(783).

Vaz, J., Jr. and Rodrigues, W. A., Jr. (1995). In *The Theory of the Electron*, *Vol. 7*(S) J. Keller and Z. Oziewicz, eds., *Advances in Applied Clifford Algebras*, Universidad Nacional Autónoma de México, pp. 369–385; http://www.arXiv.org/abs/hep-th/9511181