PHOTON AS A QUANTUM PARTICLE *

IWO BIAŁYNICKI-BIRULA

Center for Theoretical Physics, Polish Academy of Sciences
Lotników 32/46, 02-668 Warsaw, Poland
and
Institute of Theoretical Physics, Warsaw University
Hoża 69, 00-681 Warsaw, Poland
birula@cft.edu.pl

(Received January 17, 2006)

Our present understanding of the nature of photons significantly differs from what has been known years ago when the concept of a photon has only been emerging. Unfortunately, very little of this knowledge trickles to those students who do not specialize in theoretical physics. In this lecture, in addition to giving a historical perspective on the "problem of the photon", I shall say something about the description of the photon as a quantum mechanical particle. In addition, I shall show how the quantum description merges with the classical description of the electromagnetic field.

PACS numbers: 03.65.–w, 03.50.De, 42.50.–p

This paper is an expanded version of my lecture published in Polish in Postępy Fizyki. It has two parts. In the historical part I present, with very brief commentary, the views of prominent physicists of the early era of quantum theory on the nature of photons. In the contemporary part, I present my own views on the subject: How to describe the photons today. I shall show that the most natural and convenient tool is a complex combination of the electric and magnetic field vectors — the Riemann–Silberstein vector. The Riemann–Silberstein vector on the one hand contains full information about the state of the classical electromagnetic field and on the other hand it may serve as the photon wave function in the quantum theory.

* Presented at the PHOTON2005 Conference, 31 August–4 September 2005, Warsaw, Poland.
1. Photon — historical snapshots

1.1. Planck and Einstein

The notion of the photon (but not the name), as a quantum of electromagnetic radiation, made its appearance hundred years ago in the paper by Albert Einstein [1] entitled “On a heuristic point of view about the creation and conversion of light”. The key words in this paper read (English translation of the German original is taken from [2]):

“According to the assumption considered here, when a light ray starting from a point is propagated, the energy is not continuously distributed over an ever increasing volume, but it consists of a finite number of energy quanta, localized in space, which move without being divided and which can be absorbed or emitted only as a whole.”

In this paper Einstein wrote the now-famous equation connecting the kinetic energy of the electron hitting the metallic surface with the frequency of the emitted quantum of radiation. The present day version of this equation

\[ eV = h\nu - P \]

differs from the original one

\[ \Pi \epsilon = (R/N)\beta\nu - P \]

that contained the Planck constant only in a camouflaged form. From our contemporary shortened perspective it may seem that almost no time was needed to widely accept the Einstein proposal. In reality, this process took almost twenty years! This is what Max Planck had to say about the Einstein photon hypothesis in his Nobel lecture “The genesis and present state of development of the quantum theory” delivered on June 2 1920 (he received the prize in 1918 but due to perturbations at the end of the WWII, he went to Stockholm two years later):

“There is in particular one problem whose exhaustive solution could provide considerable elucidation. What becomes of the energy of a photon after complete emission? Does it spread out in all directions with further propagation in the sense of Huyghens’ wave theory, so constantly taking up more space, in the boundless progressive attenuation? Or does it fly out like a projectile in one direction in the sense of Newton’s emanation theory? In the first case, the quantum would no longer be in the position to concentrate energy upon a single point in space in such a way as to release an electron from its atomic bond, and in the second case, the main triumph of the Maxwell theory — the continuity between the static and the dynamic fields and, with it, the complete understanding we have enjoyed, until now, of the fully investigated interference phenomena — would have to be sacrificed, both being very unhappy consequences for today’s theoreticians.”
He finished his lecture on a more optimistic note:

“Be that as it may, in any case no doubt can arise that science will master the dilemma, serious as it is, and that which appears today so unsatisfactory will in fact eventually, seen from a higher vintage point, be distinguished by its special harmony and simplicity. Until this aim is achieved, the problem of the quantum of action will not cease to inspire research and fructify it, and the greater the difficulties which oppose its solution, the more significant it finally will show itself to be for the broadening and deepening of our whole knowledge in physics.”

The translation of the lecture is taken from the web site of the Nobel Committee. It must be noted that this translation was made without proper care for historical correctness. In the Nobel lecture delivered in German and published later as a booklet [3] Planck refers always to “Lichtquantum” (light quantum). The careless translator writes “photon” apparently without realizing that the term photon was introduced six years after the Nobel lecture by Planck. It is a pity that the Nobel Committee did not use a translation published in 1922 [4].

The slow acceptance of the light quanta is illustrated by the fact that Planck and Einstein received the Nobel prizes eighteen and sixteen years after their discoveries. Planck received the prize “In recognition of the services he rendered to the advancement of Physics by his discovery of energy quanta” and Einstein received not for his discovery of the photon but “For his services to Theoretical Physics, and especially for his discovery of the law of the photoelectric effect”.

1.2. Millikan and Compton

There is no better illustration of the closing words of Planck’s lecture that “quantum of action will inspire research” than relentless, lasting many years experiments of Robert Millikan on the photoelectric effect. Millikan is mostly remembered for his measurement of the elementary charge, but his Nobel prize was awarded for two achievements: “for his work on the elementary charge of electricity and on the photoelectric effect”. His Nobel lecture delivered May 23 1924, almost twenty years after Einstein’s paper, had the title: The electron and the light-quant(sic!) from the experimental point of view. In this lecture he gave a summary of his work but also he included a very apt summary of the role of the experiment and the theory in physics. I could not resist the temptation to quote this part of his talk because today the physicists should even more take it to heart than eighty years ago:

“The fact that Science walks forward on two feet, namely theory and experiment, is nowhere better illustrated than in the two fields for slight
contributions to which you have done me the great honor of awarding me the Nobel Prize in Physics for the year 1923. Sometimes it is one foot which is put forward first, sometimes the other, but continuous progress is only made by the use of both — by theorizing and then testing, or by finding new relations in the process of experimenting and then bringing the theoretical foot up and pushing it on beyond, and so on in unending alternations.”

This is how Millikan describes the laborious and long road to a final experimental test of the Einstein hypothesis:

“After ten years of testing and changing and learning and sometimes blundering, all efforts being directed from the first toward the accurate experimental measurement of the energies of emission of photoelectrons, now as a function of temperature, now of wavelength, now of material (contact e. m. f. relations), this work resulted, contrary to my own expectation, in the first direct experimental proof in 1914 of the exact validity, within narrow limits of experimental error, of the Einstein equation, and the first direct photoelectric determination of Planck’s h.”

The Einstein theory of the photoelectric effect finally triumphed:

“This work, like that on the electron, has had to run the gauntlet of severe criticism, for up to 1916 not only was discussion active as to whether there were any limiting velocity of emission, but other observers who had thought that a linear relation existed between energy and frequency had not found the invariable constant h appearing as the ratio. But at the present time it is not too much to say, that the altogether overwhelming proof furnished by the experiments of many different observers, working by different methods in many different laboratories, that Einstein’s equation is one of exact validity (always within the present small limits of experimental error) and of very general applicability, is perhaps the most conspicuous achievement of Experimental Physics during the past decade.”

However, there were still doubts about the reality of light quanta:

“In view of all these methods and experiments the general validity of Einstein’s equation is, I think, now universally conceded, and to that extent the reality of Einstein’s light-quanta may be considered as experimentally established. But the conception of localized light-quanta out of which Einstein got his equation must still be regarded as far from being established.”

The final confirmation came from experiments with the quanta of higher energy — with X-ray photons. Millikan devoted a part of his lecture to these experiments:

“Within the past year, however, a young American physicist, Arthur H. Compton of the University of Chicago, by using the conception of localized light-quanta, has brought forward another new phenomenon which at least shows the fecundity of the Einstein hypothesis. Compton goes a step farther than Einstein in that he assumes not only the existence of light-
quanta but also that in the impact between a light-quant and a free electron the laws of conservation of energy and of conservation of momentum both hold. This assumption enables him to compute exactly how much the frequency of ether waves which have collided with free electrons will be lowered because of the energy which they have given up to the electron in the act of collision, and therefore the loss which their own $h\nu$ has experienced. He then finds experimentally that there is approximately the computed lowering in frequency when monochromatic X-rays from molybdenum are scattered by carbon."

And this is how Arthur Compton described his discovery [6]:

“The present theory depends essentially upon the assumption that each electron which is effective in the scattering scatters a complete quantum. It involves also the hypothesis that the quanta of radiation are received from definite directions and are scattered in definite directions. The experimental support of the theory indicates very convincingly that a radiation quantum carries with it directed momentum as well as energy. Emphasis has been laid upon the fact that in its present form the quantum theory of scattering applies only to light elements. The reason for this restriction is that we have tacitly assumed that there are no forces of constraint acting upon the scattering electron.”

It should be mentioned here that Millikan was not right when he said that Compton went “a step further than Einstein” because seven years before Compton announced his results Einstein published the now-famous paper [5] in which he laid foundations of the laser theory. In this paper we read: “If a ray of light causes molecules hit by it to absorb or emit through an elementary process an amount of energy $h\nu$ in the form of radiation (induced radiation process), the momentum $h\nu/c$ is always transferred to the molecule, and in such a way that the momentum is directed along the direction of propagation of the ray if the energy is absorbed, and directed in the opposite direction, if the energy is emitted.” Compton was probably unaware of the Einstein paper. As a matter of fact, he does not refer to any theoretical work on photons in his paper [6] but that might have been due to a different habits in those days.

Thus, all seemed to indicate that the photons exist but Millikan (and many others) was still pondering on the wave-particle duality:

“It may be said then without hesitation that it is not merely the Einstein equation which is having extraordinary success at the moment, but the Einstein conception as well. But until it can account for the facts of interference and the other effects which have seemed thus far to be irreconcilable with it, we must withhold our full assent.”
1.3. Bohr’s mistake

Similar reservations were expressed by Niels Bohr. In the paper [7] that he published with his two young collaborators we find the following sentence:

“Although the great heuristic value of this hypothesis is shown by the confirmation of Einstein’s predictions concerning the photoelectric phenomenon, still the theory of light quanta can obviously not be considered as a satisfactory solution of the problem of light propagation.”

This paper contained the most unfortunate idea of Bohr. To reconcile the corpuscular and the wave properties of light, the authors put forward a hypothesis that the energy conservation in processes with the participation of photons holds only on average — for the mean values. Soon afterwards, however, Walter Bothe and Hans Geiger proved experimentally, by measuring the coincidences between the emitted photons and recoiled electrons that the conservation laws hold in each individual act of the photon collision with an electron. Bothe will receive the Nobel prize for this work in 1954.

1.4. Photon is named

Twenty years have passed after the publication of the Einstein paper and the term “photon” still had not been invented. It made its appearance in a letter to Nature [8] written by Gilbert N. Lewis, professor of theoretical chemistry in Berkeley. In this letter he wrote:

“Had there not seemed to be insuperable objections, one might have been tempted to adopt the hypothesis that we are dealing here with a new type of atom, an identifiable entity, uncreatable and indestructible, which acts as the carrier of radiant energy and, after absorption, persists as an essential constituent of the absorbing atom until it is later sent out again bearing a new amount of energy.”

and:

“It would seem inappropriate to speak of one of these hypothetical entities as a particle of light, a corpuscle of light, a light quantum, or a light quant, if we are to assume that it spends only a minute fraction of its existence as a carrier of radiant energy, while the rest of the time it remains as an important structural element within the atom. It would also cause confusion to call it merely a quantum, for later it will be necessary to distinguish between the number of these entities present in an atom and the so-called quantum number. I therefore take the liberty of proposing for this hypothetical new atom, which is not light but plays an essential part in every process of radiation, the name photon.”

A year later Paul A. M. Dirac published his quantum theory of the emission and absorption of radiation [9] which led to a full unification of the corpuscular and the wave points of view. This unification became possible
due to the emergence of quantum mechanics. Quantum theory of electromagnetic radiation started by Dirac and elaborated further by Heisenberg and Pauli [10] and Fermi [11] was the cornerstone of quantum electrodynamics — a full relativistic theory of charged particles in interaction with the electromagnetic field. Paradoxically, despite of the existence of a complete quantum theory of electromagnetic radiation for almost eighty years now, this knowledge has not trickled down to university textbooks. Nearly all of them present the quantum theory of the photon of the ancient era. This is so even in basic textbooks devoted to quantum theory. The only exception is Fundamentals of Physics by Halliday, Resnick, and Walker [12]. In the volume 5 of this very popular textbook we find an almost correct description of the connection between the corpuscular and wave properties of photons.

"The probability (per unit time interval) that a photon will be detected in any small volume centered on a given point in a light wave is proportional to the square of the amplitude of the wave’s electric field vector at that point. We now have a probabilistic description of a light wave, hence another way to view light. It is not only an electromagnetic wave but also a probability wave. That is to every point in a light wave we can attach a numerical probability (per unit time interval) that a photon can be detected in any small volume centered on that point."

Of course, we still do not understand the nature of photons to the very end, like we do not understand the quantum theory, which should not keep us from describing photons as quantum particles. It is true that in contrast to massive particles, massless particles demand more sophisticated methods of description. In this case, the nonrelativistic wave mechanics with the traditional Schrödinger equation does not exist. We may even say that the following words of Einstein found in a letter of December 12, 1951 to his friend Michele Angelo Besso are still valid:

“All the fifty years of conscious brooding have brought me no closer to the answer of the question: What are light quanta? Of course, today every rascal thinks he knows the answer, but he is deluding himself.”

2. Photons today

In nature there are two kinds of photons: right-handed and left-handed. For a right-handed/left-handed photon, the projection of its angular momentum on the direction of its momentum is positive/negative. According to the group-theoretic classification, these are two different elementary particles — they are described by two distinct irreducible representation of the Poincaré group. In plain language this means that by shifting, rotating, and boosting to a moving frame one cannot transform a righthanded into a lefthanded one. This is quite different from the behavior of spinning mas-
sive particles, for example, electrons. An electron with a positive projection of its angular momentum on its momentum becomes an electron with the negative projection when boosted to a moving frame moving in which the electron velocity is reversed.

2.1. Quantum mechanics of photons

Photons are, without any doubt, quantum particles — their behavior is governed by the laws of quantum mechanics. This means, among other things, that their (pure) state are described by wave functions. Quantum-mechanical wave functions obey the principle of superposition — a sum of two wave functions describes an allowed state of the photon. However, one cannot add the wave functions of two different photons because the right-handed and left-handed photons are described by two different representations of the Poincaré group. Similarly, as is well known, one cannot add directly two different components of a vector, say $x$ and $y$ components. How does one then form a superposition of photon wave functions to produce states with linear polarization, or in general, with an elliptical polarization? We can do it exactly as in the case of ordinary vectors in a plane. Namely, one needs the analogues of the unit vectors that appear in the composition law for vectors

\[ \mathbf{a} = a_x \mathbf{i}_x + a_y \mathbf{i}_y. \] (3)

In the case of photons, one can write

\[ |\psi\rangle = f_+ (k) |R\rangle + f_- (k) |L\rangle, \] (4)

where the scalar product in the space of states and the normalization condition, that are needed for the probabilistic interpretation, have the form

\[ \langle \phi | \psi \rangle = \int \frac{d^3 k}{(2 \pi)^3 \omega} \left( g_+^* (k) f_+ (k) + g_-^* (k) f_- (k) \right), \] (5)

\[ \langle \psi | \psi \rangle = \int \frac{d^3 k}{(2 \pi)^3 \omega} \left( |f_+ (k)|^2 + |f_- (k)|^2 \right) = 1. \] (6)

The appearance of the frequency $\omega$ in the denominator is characteristic of relativistic theories. Since the ratio $d^3 k/\omega$ is invariant under Lorentz transformation, the product $g_\pm^*(k) f_\pm (k)$ must be a scalar — the functions only change their phases under all transformations. The functions $f_\pm (k)$ have the interpretation of probability amplitudes in momentum representation. This means that $|f_\pm (k)|^2/(2 \pi)^3 \omega$ is the probability density (per unit volume in the space of wave vectors) to find the right-handed/left-handed photon with the momentum $\hbar k$. 
Photon as a Quantum Particle

Exactly, as in quantum mechanics of massive particles, the physical variables in quantum mechanics of photons are represented by operators. We can define these operators according to the rules of the standard quantum mechanics in momentum representation. The energy and momentum are simply represented by the appropriate multiplication operators and the angular momentum operator involves also differentiation

\[
\hat{E} = \hbar \omega = \hbar c |k|, \\
\hat{p} = \hbar k, \\
\hat{M}_z = -i \hbar (k_x \partial_{k_y} - k_y \partial_{k_x}).
\] (7)

\[
\hat{E} = \hbar \omega = \hbar c |k|, \\
\hat{p} = \hbar k, \\
\hat{M}_z = -i \hbar (k_x \partial_{k_y} - k_y \partial_{k_x}).
\] (8)

\[
\hat{E} = \hbar \omega = \hbar c |k|, \\
\hat{p} = \hbar k, \\
\hat{M}_z = -i \hbar (k_x \partial_{k_y} - k_y \partial_{k_x}).
\] (9)

2.2. Riemann–Silberstein vector — photon wave function

The remaining task is to find the connection between the photon wave functions \( f_{\pm}(k) \) and the classical electromagnetic field. This is done in a natural way with the use of the Riemann–Silberstein (RS) vector \( \mathbf{F}(r, t) \)

\[
\mathbf{F}(r, t) = \sqrt{\frac{\varepsilon_0}{2}} (\mathbf{E}(r, t) + ic\mathbf{B}(r, t)).
\] (10)

This complex vector carries exactly the same information as two real vectors \( \mathbf{E} \) and \( \mathbf{B} \). The Maxwell equations expressed as equations for the RS vector read

\[
i \partial_t \mathbf{F} = c \mathbf{\nabla} \times \mathbf{F}, \quad \mathbf{\nabla} \cdot \mathbf{F} = 0.
\] (11)

After the multiplication by \( \hbar \), the time evolution equation for \( \mathbf{F} \) can be written in the form of Schrödinger equation

\[
i \hbar \partial_t \mathbf{F} = c (\hat{\mathbf{p}} \cdot \hat{\mathbf{s}}) \mathbf{F},
\] (12)

where I replaced the nabla by the quantum-mechanical momentum operator and the curl by a multiplication by the following spin one matrices

\[
s_x = \begin{pmatrix}
0 & 0 & 0 \\
0 & 0 & -i \\
i & 0 & 0
\end{pmatrix},
\quad
s_y = \begin{pmatrix}
0 & 0 & i \\
0 & 0 & 0 \\
-i & 0 & 0
\end{pmatrix},
\quad
s_z = \begin{pmatrix}
0 & -i & 0 \\
i & 0 & 0 \\
0 & 0 & 0
\end{pmatrix}.
\] (13)

Classical electromagnetic field is made of a huge number of photons — all of them in the same quantum state. This field can be directly connected with the photon wave functions with the use of the RS vector. The easiest
way to proceed is to expand an arbitrary solution of the Maxwell equations for $\mathbf{F}$ into plane waves according to the formula

$$
\mathbf{F}(\mathbf{r},t) = \int d^3k \; \mathbf{e}(\mathbf{k}) \sqrt{\langle N \rangle/\hbar \omega} \left( f_+(\mathbf{k}) e^{-i\omega t+i\mathbf{k} \cdot \mathbf{r}} + f_-^*(\mathbf{k}) e^{i\omega t-ik \cdot r} \right), \quad (14)
$$

where $\langle N \rangle$ denotes the mean number of photons in the wave. The complex vector $\mathbf{e}(\mathbf{k})$ is a solution of the algebraic equation

$$
\mathbf{k} \times \mathbf{e}(\mathbf{k}) = -i\mathbf{k} \mathbf{e}(\mathbf{k}) \quad (15)
$$

normalized to one. One may be puzzled, why in the formula (14) a natural symmetry between $f_+$ and $f_-$ has been destroyed? This is due to our choice of convention. Since I have chosen $\mathbf{E} + ic\mathbf{B}$ as the fundamental quantity, instead of $\mathbf{E} - ic\mathbf{B}$, I have broken the symmetry between the right-handed and the left-handed photons.

The expansion (14) of the RS vector shows that the quantum wave functions of the photons in momentum space play also the role of the amplitudes in the expansion of the classical electromagnetic field into plane waves. The precise connection between the probabilities to detect a photon, as described by the wave functions $f_+$ and $f_-$, and the classical electric field (the real part of $\mathbf{F}$) is more complicated than what one finds in the HRW textbook [12]. These complications arise because in the case of photons the exact analog of the wave function in the position representation does not exist. The RS vector, treated as a photon wave function, does carry all the information about the state of the photon, because we can recover from it the wave functions $f_+$ and $f_-$ in momentum representation. However, the RS vector does not possess the standard probabilistic interpretation. The modulus squared of this vector $\mathbf{F}^* \cdot \mathbf{F}$ has the dimension of the energy density and indeed it is the energy density in the classical theory. Therefore, the quantity

$$
\int_V d^3x \mathbf{F}^* \cdot \mathbf{F} \quad (16)
$$

can only be used as a measure of the average energy of the photon in a given volume $V$. There is an obvious difficulty in extracting the probability of finding a photon in a given volume from the average energy. In order to find out what is the energy distribution from the knowledge of $\mathbf{F}$ one must perform a Fourier decomposition to determine the amplitudes $f_+$ and $f_-$. The Fourier analysis is a nonlocal operation, since it involves integration over all space. Thus, one cannot repeat the treatment known from the Schrödinger wave mechanics and define the (local) probability density to
find a photon in a vicinity of some point $r$. Nevertheless, despite the lack of the standard probabilistic interpretation in position space, one may use the RS vector as a substitute for the wave function. After all, this vector characterizes the state of the photon in a complete fashion, it enables one to find many important characteristics of the photon state (energy distribution, momentum and angular momentum distribution, etc.), and it gives one a tool to formulate the quantum superposition principle (one may form a new state by taking a sum of any two states with arbitrary complex coefficients).

The quantum theory of the electromagnetic field employing the photon wave functions in the form of the RS vector requires at some point the introduction of the operators of the electromagnetic field. This does not mean, however, that the photon wave function looses its significance. Even in the field-theoretic approach based on the creation and annihilation operators one needs the photon wave functions to distinguish different photon states — one needs to know precisely which particular photon states are being created and annihilated. In other words, in full-fledged quantum electrodynamics the photon wave functions serve as labels — they label the creation and annihilation operators. Of course, there is a complete agreement between the standard formulation of the quantum theory of the electromagnetic field and the one based on the photon wave function. The use of photon wave functions enriches the theory by bridging a gap, existing in the standard formulation, between the classical theory and the quantum theory — it provides the counterpart of the first quantization. In particular, it is a very useful tool in the description of light beams [16] where the classical wave description is intermingled with the photon description.

2.3. Quantum electrodynamics

The special role played by the RS vector in the quantum mechanics of photons, carries over to quantum field theory of the electromagnetic field. In this theory, the electric and magnetic field vectors become field operators and so does the RS vector. Instead of the decomposition into the plane waves (14), one now has the operator formula

$$\vec{F}(\vec{r}, t) = \int d^3k \, e(k) \sqrt{\frac{\hbar \omega}{(2\pi)^3}} \left( \hat{a}_+(k)e^{-i\omega t + i\vec{k} \cdot \vec{r}} + \hat{b}^\dagger_-(k)e^{i\omega t - i\vec{k} \cdot \vec{r}} \right),$$

with the annihilation and creation operators replacing the classical amplitudes. This expression is much simpler than the expansion of the electric and magnetic field because there is only one term for the right-handed and the left-handed photons. As a consequence, the operators $\vec{F}$ at different space-time points commute and, therefore, in a quantum state of the electromagnetic field the values of the RS vector at different points can be known
precisely — there are no uncertainty relations to satisfy. Such a state is known as the coherent state. It is the closest counterpart of the classical state characterized by the same RS vector. Of course, the formulation of quantum electrodynamics employing the RS vector is in full agreement with the standard approach but certain characteristic properties of the electromagnetic field come out much more clearly.

REFERENCES

[5] This paper is best known from its version that appeared a year later: A. Einstein Zur Quantenlehre der Strahlung, Physikalische Zeitschrift 18, 121 (1917); English translation can be found in Ref [2].