EINSTEIN’S REVOLUTIONARY LIGHT–QUANTUM HYPOTHESIS*

ROGER H. STUEWER

School of Physics and Astronomy, University of Minnesota
116 Church St. SE, Minneapolis, MN 55455, USA

(Received December 6, 2005)

Albert Einstein’s light-quantum paper was the only one of his great papers of 1905 that he himself called “very revolutionary”. I sketch his arguments for light quanta, his analysis of the photoelectric effect, and his introduction of the wave-particle duality into physics in 1909. I show that Robert Andrews Millikan, in common with almost all physicists at the time, rejected Einstein’s light-quantum hypothesis as an interpretation of his photoelectric-effect experiments of 1915. I then trace the complex experimental and theoretical route that Arthur Holly Compton followed between 1916 and 1922 that led to his discovery of the Compton effect, a discovery that Peter Debye also made virtually simultaneously and independently. Compton’s discovery, however, was challenged on experimental grounds by William Duane and on theoretical grounds by Niels Bohr in the Bohr–Kramers–Slater theory of 1924, and only after that theory was disproved experimentally the following year by Walther Bothe and Hans Geiger in Berlin and by Compton and Alfred W. Simon in Chicago was Einstein’s light-quantum hypothesis generally accepted by physicists.

PACS numbers: 01.65.+g

Albert Einstein signed his light-quantum paper, “Concerning a Heuristic Point of View about the Creation and Transformation of Light” [1], in Bern, Switzerland, on March 17, 1905, three days after his twenty-sixth birthday. It was the only one of his great papers of 1905 that he himself called “very revolutionary” [2]. As we shall see, Einstein was correct: His light-quantum hypothesis was not generally accepted by physicists for another two decades.

Einstein gave two arguments for light quanta, a negative and a positive one [3]. His negative argument was the failure of the classical equipartition

* Presented at The Centenary of the Photon Conference, 30 August 2005, Warsaw, Poland.
theorem — what Paul Ehrenfest in 1911 called the “ultraviolet catastrophe” — which showed that there was something fundamentally wrong with classical radiation theory. His positive argument proceeded in two stages. First, he calculated the change in entropy when a volume $V_0$ filled with blackbody radiation of total energy $U$ in the Wien’s law (high-frequency) region of the spectrum was reduced to a subvolume $V$. Second, he used Ludwig Boltzmann’s statistical version of the second law of thermodynamics to calculate the probability of finding $n$ independently moving, distinguishable gas molecules in a volume $V_0$ at a given instant of time in a subvolume $V$. These two results were formally identical, providing that

$$U = n \left( \frac{R\beta}{N} \right) \nu,$$

where $R$ is the ideal gas constant, $\beta$ is the constant that appears in the exponent in Wien’s law, $N$ is Avogadro’s number, and $\nu$ is the frequency of the radiation. Einstein concluded:

Monochromatic radiation of low density (within the range of validity of Wien’s radiation formula) behaves thermodynamically as if it consisted of mutually independent energy quanta of magnitude $R\beta\nu/N$ [4].

Einstein cited three experimental supports for his light-quantum hypothesis, the most famous one being the photoelectric effect, which was discovered by Heinrich Hertz in 1886 [5] and explored experimentally by Philipp Lenard in 1902 [6]. Einstein wrote down his famous equation of the photoelectric effect,

$$\Pi e = \frac{R}{N} \beta \nu - P,$$

where $\Pi$ is the potential necessary to prevent the emission of an electron of charge $e$, and $P$ is the energy lost by the electron in reaching the surface of the irradiated metal, the so-called work function. It would take a decade to confirm this equation experimentally.

Einstein also pointed out, at the same time, that if the incident light-quantum did not transfer all of its energy to the electron, then the above equation would become an inequality,

$$\Pi e < \frac{R}{N} \beta \nu - P,$$

which would take almost two decades to confirm experimentally.

We therefore see that Einstein’s positive argument for light quanta was based upon the second law of thermodynamics in its statistical interpretation. He did not propose his light-quantum hypothesis “to explain the
Einstein’s Revolutionary Light–Quantum Hypothesis

photoelectric effect”. That was only one of three experimental supports he cited for it, so to call Einstein’s paper his “photoelectric-effect paper” is completely false historically and utterly trivializes Einstein’s achievement.

In January 1909 Einstein presented a further argument for light quanta based upon his analysis of the energy and momentum fluctuations in blackbody radiation [7]. He now assumed the validity of Planck’s law and showed that the expressions for the mean-square energy and momentum fluctuations split naturally into a sum of two terms, a wave term that dominated in the Rayleigh–Jeans (low-frequency) region of the blackbody spectrum, and a particle term that dominated in the Wien’s law (high-frequency) region, so that both terms were necessary to describe the fluctuations for the complete blackbody spectrum. This constituted Einstein’s introduction of the wave-particle duality into physics [8]. He presented these ideas again at a meeting in Salzburg, Austria, in September 1909 [9]. During the discussion, Max Planck took the acceptance of Einstein’s light quanta to imply the rejection of Maxwell’s electromagnetic waves, which, he said, “seems to me to be a step which in my opinion is not yet necessary” [10]. The only physicist who supported Einstein’s light-quantum hypothesis was Johannes Stark.

In general, by around 1913 almost all physicists rejected Einstein’s light-quantum hypothesis, and they had good reasons for doing so [11]. First, they believed that Maxwell’s electromagnetic theory had to be universally valid to account for interference and diffraction phenomena. Second, Einstein’s statistical arguments for light quanta were unfamiliar and difficult to grasp. Third, between 1910 and 1913 three prominent physicists, J.J. Thomson, Arnold Sommerfeld, and O.W. Richardson, showed that Einstein’s equation of the photoelectric effect could be derived on classical, non-Einsteinian grounds, thereby obviating the need to accept Einstein’s light-quantum hypothesis as an interpretation of it. Fourth, in 1912 Max Laue, Walter Friedrich, and Paul Knipping showed that X-rays can be diffracted by a copper-sulfate crystal, which everyone took to be clear proof that they were electromagnetic waves of short wavelength. Fifth, in 1913 Niels Bohr insisted that when an electron undergoes a transition between two stationary states in a hydrogen atom, an electromagnetic wave, not a light quantum, is emitted — a point to which I shall return later.

Robert Andrews Millikan recalled that he began carrying out his famous photoelectric-effect experiments in earnest in October 1912, which then “occupied practically all of my individual research time for the next three years” [12]. He eventually found “the key to the whole problem”, namely, that radiation over a wide range of frequencies ejected photoelectrons from the highly electropositive alkali metals, lithium, sodium, and potassium. He modified and improved his apparatus repeatedly until it became, he said, “a machine shop in vacuo” [13]. He reported his results at a meeting of the
American Physical Society in April 1915 and published them in the *Physical Review* in March 1916 [14]. They left no doubt about the complete validity of Einstein’s equation of the photoelectric effect, as seen in figure 1, where Millikan’s data points for sodium fell on a perfectly straight line of slope $h/e$.

Fig. 1. Millikan’s plot of the maximum energies of the photoelectrons emitted from sodium as a function of the frequencies of the incident radiation. His data points all fell on a straight line of slope $h/e$ as required by Einstein’s equation of the photoelectric effect. Source: Millikan, “Direct Photoelectric Determination” Ref. [14], p. 373.

That left the interpretation of his results. In Millikan’s *Autobiography*, which he published in 1950 at the age of 82, he included a chapter on “The Experimental Proof of the Existence of the Photon” in which he wrote:

This seemed to me, as it did to many others, a matter of very great importance, for it . . . proved simply and irrefutably I thought, *that the emitted electron that escapes with the energy $hv$ gets that energy by the direct transfer of $hv$ units of energy from the light to the electron* and hence scarcely permits of any other interpretation than that which Einstein had originally suggested, namely that of the semi-corpuscular or photon theory of light itself [Millikan’s italics] [15].

In Millikan’s paper of 1916, however, which he published at the age of 48, we find a very different interpretation. Millikan declared there that Einstein’s “bold, not to say reckless” light-quantum hypothesis “flies in the face of the
thoroughly established facts of interference” [16], so that we must search for “a substitute for Einstein’s theory”. Millikan’s “substitute” theory was that the photosensitive metal must contain “oscillators of all frequencies” that “are at all times . . . loading up to the [energy] value $h\nu$” [17]. A few of them will be “in tune” with the frequency $\nu_0$ of the incident radiation and thus will absorb energy until it reaches the “critical value” $h\nu_0$, at which time an “explosion” will occur and the electron will be “shot out” from the atom.

Millikan therefore fell completely in line with Thomson, Sommerfeld, and Richardson in proposing a classical, non-Einsteinian theory of the photoelectric effect in his paper of 1916. No one, in fact, made Millikan’s views on Einstein’s light-quantum hypothesis clearer than Millikan himself did in his book, The Electron, which he published in 1917, where he wrote:

> Despite . . . the apparently complete success of the Einstein equation, the physical theory of which it was designed to be the symbolic expression is found so untenable that Einstein himself, I believe, no longer holds to it, and we are in the position of having built a very perfect structure and then knocked out entirely the underpinning without causing the building to fall. It [Einstein’s equation] stands complete and apparently well tested, but without any visible means of support. These supports must obviously exist, and the most fascinating problem of modern physics is to find them. Experiment has outrun theory, or, better, *guided by erroneous theory*, it has discovered relationships which seem to be of the greatest interest and importance, but the reasons for them are as yet not at all understood [18].

This, note, is the same man who thirty-four years later, in 1950, wrote that his experiments “proved simply and irrefutably I thought”, that they scarcely permit “any other interpretation than that which Einstein had originally suggested, namely that of the semi-corpuscular or photon theory of light”.

Historians have a name for this, namely, “revisionist history”. This, however, is by no means the only time that Millikan revised history. The earliest instance I found dates to 1906 when he and Henry G. Gale, his younger colleague at the University of Chicago, published their textbook, *A First Course in Physics* [19]. In it Millikan reproduced a picture taken in 1899 of J.J. Thomson reading a newspaper in his study while sitting in a chair once owned by James Clerk Maxwell [20], which was identical to the original picture — except that Millikan carefully etched out the cigarette in Thomson’s left hand, as shown in figure 2, presumably because he did not wish to corrupt young and impressionable students at Chicago and elsewhere. This and his 1950 interpretation of his photoelectric-effect experiments illustrate what I like to call Millikan’s philosophy of history: “If the facts don’t fit your theory, change the facts”.
Millikan’s rejection of Einstein’s light-quantum hypothesis in 1916–1917, however, was completely in line with the almost universal attitude of physicists toward it at that time — which was the atmosphere in which Arthur Holly Compton began his work [21]. Compton received his Ph.D. degree in physics at Princeton University in 1916 and then spent one year at the University of Minnesota (1916–1917), two years at the Westinghouse Electric and Manufacturing Company in Pittsburgh (1917–1919), and one year as a National Research Council Fellow at the Cavendish Laboratory in Cambridge (1919–1920) before being appointed as Wayman Crow Professor and Head of the Department of Physics at Washington University in St. Louis in 1920, where he remained until he moved to the University of Chicago in 1923.

While Compton was at Westinghouse, he came across a puzzling observation that C.G. Barkla had made in 1917, namely, that the mass-absorption coefficient for 0.145 Ångstrom X-rays in aluminum was markedly smaller than the Thomson mass-scattering coefficient — which seemed impossible, because the mass-absorption coefficient consists of the sum of the mass-fluorescent and mass-scattering coefficients. In other words, how could the total be less than one of its parts? Compton thought long and hard about
this apparent contradiction and eventually came up with an explanation of it: He concluded that the incident X-rays were being diffracted by the electrons in aluminum. For diffraction to occur, however, the diameter of the diffracting obstacle, the electron, had to be on the order of the wavelength of the incident X-rays, about 0.1 Ångstrom — in other words, it had to be almost as large as the Bohr radius of the hydrogen atom, a very large electron indeed. That was too much for Ernest Rutherford, who when Compton was at the Cavendish Laboratory invited Compton to give a talk at a meeting of the Cambridge Philosophical Society and introduced him as follows: “This is Dr. Compton who is here to talk to us about the Size of the Electron. Please listen to him attentively, but you don’t have to believe him” [22]. Charles D. Ellis, who was in Compton’s audience, recalled that at one point Rutherford burst out saying, “I will not have an electron as big as a balloon in my Laboratory” [23].

In addition to hearing such friendly criticism, Compton carried out gamma-ray scattering experiments at the Cavendish Laboratory that would contribute greatly to the further evolution of his thought. He found in his experiments, first, that the intensity of the scattered gamma rays was greater in the forward than in the backward direction; second, that the scattered gamma rays were “softer” or of greater wavelength than the primary gamma rays; third, that the “hardness” or wavelength of the scattered gamma rays was independent of the nature of the scatterer; and fourth, that the scattered gamma rays became “softer” or of greater wavelength as the scattering angle increased. We of course recognize these results as characteristic of the Compton effect, but the question is: How did Compton himself interpret them in 1919–1920? Once again Compton thought long and hard about them, and since he too was completely convinced that X-rays were electromagnetic waves of short wavelength, he concluded that the incident gamma rays were exciting the emission of a new type of longer-wavelength “fluorescent” gamma radiation in the scatterer: The incident gamma rays were sending the electrons in the scatterer, which he viewed as tiny electron-oscillators, forward at high velocities as they emitted this secondary “fluorescent” radiation. This radiation is peaked in the forward direction, which explained the forward–backward asymmetry in its intensity, while the Doppler shift explained its increased wavelength.

That was the interpretation that Compton had in mind when he left the Cavendish Laboratory and arrived at Washington University in St. Louis in the summer of 1920. He took a Bragg spectrometer along with him from the Cavendish, because he wanted to carry out similar X-ray scattering experiments at Washington University. He reported his first results in December 1921. He sent monochromatic MoKα X-rays (wavelength 0.708 Ångstrom) onto a pyrex scatterer and viewed the secondary spectrum at a scattering
angle of about 90°, as shown in figure 3. I emphasize that these are my plots of Compton’s experimental data as recorded in his laboratory notebooks, because I knew what I was looking for, but Compton did not. Thus, Compton’s published paper leaves no doubt that he did not see the small change in wavelength between the primary and secondary spectra as I have shown it, but instead reported that the wavelength of the scattered X-rays was 35% greater than that of the primary X-rays, or 0.95 Ångstrom. In other words, he took the secondary spectrum to consist of the low peaks on the right (which we recognize as simply the second-order spectrum). To Compton, the ratio of the wavelength \( \lambda \) of the primary X-rays to the wavelength \( \lambda' \) of the scattered X-rays was \( \lambda/\lambda' = (0.708 \, \text{Å})/(0.95 \, \text{Å}) = 0.75. \)

![Fig. 3. My plots of Compton’s spectra of December 1921 based upon data recorded in his laboratory notebooks. He sent monochromatic MoK\(_\alpha\) X-rays onto a pyrex scatterer and observed the secondary spectrum at a scattering angle of about 90°. The relative intensity of the primary and scattered MoK\(_\alpha\) X-rays is plotted against their glancing angle \( \theta \) from the spectrometer crystal and the corresponding wavelength \( \lambda. \) Source: Stuewer, Compton Effect (Ref. [11]), p. 187.](image-url)

How then did Compton interpret this large change in wavelength? Answer: By the Doppler shift. Thus, at 90°, the Doppler shift is given by \( \lambda/\lambda' = 1 - v/c, \) where \( c \) is the speed of light and where the speed \( v \) of the electron-oscillators (mass \( m \)) can be eliminated by “conservation of energy”, namely, \( (1/2)mv^2 = h\nu, \) so that substituting we have \( \lambda/\lambda' = 1 - v/c = \)
Who could ask for better agreement between theory and experiment? I think this is a wonderful historical example of a false theory being confirmed by spurious experimental data.

By October 1922 Compton realized that he had made a mistake [24]: The increase in wavelength between the primary and secondary X-ray spectra was not 35% but only a few percent, as shown in figure 4, where he now used graphite (carbon) as the scatterer. Thus, as indicated in his annotations to the right of the spectra, he now found that the ratio of the wavelength $\lambda$ of the primary X-rays to the wavelength $\lambda'$ of the scattered X-rays was $\lambda/\lambda' = (0.708 \text{ Å})/(0.730 \text{ Å}) = 0.969$. How then did Compton interpret this small change in wavelength? Answer: Again by the Doppler shift. Thus, again at 90°, the Doppler shift is given by $\lambda/\lambda' = 1 - v/c$, where the speed $v$ of the electron-oscillators now can be eliminated by “conservation of momentum”, namely, $mv = h/\lambda$, so that substituting we have $\lambda/\lambda' = 1 - v/c = 1 - h/mc\lambda = 1 - hv/mc^2 = 1 - (0.017 \text{ MeV})/(0.51 \text{ MeV}) = 1 - 0.034 = 0.966$, which is precisely the result Compton gives to the right.

Fig. 4. Compton’s spectra of October 1922. As indicated in the annotations on the right, the change in wavelength of the MoKα X-rays when scattered by graphite (carbon) at a scattering angle of 90° was 0.730 Å – 0.708 Å = 0.022 Å, or about 3%. Source: Compton, “Secondary Radiations” (Ref. [24]), p. 16; 336.
of his spectra. Again, who could ask for better agreement between theory and experiment? I think this is a wonderful historical example of a false theory being confirmed by good experimental data.

Compton put everything together one month later, in November 1922. He now assumed that a single X-ray quantum interacts with a single electron in the scatterer in a billiard-ball collision process in which both energy and momentum are conserved. In his published paper [25], he drew his famous vector diagram of the collision process (figure 5), using the correct relativistic expression for the mass $m$ of the electron, and calculated the change in wavelength $\Delta \lambda$ (in his current notation) to be $\Delta \lambda = \lambda_0 - \lambda_0 = (h/mc)(1 - \cos \theta)$, where $\theta$ is the scattering angle. Further, to substantiate his new quantum theory of scattering, he again published primary and secondary X-ray spectra, as shown in figure 6. But note: These are exactly the same spectra that Compton had published in October 1922 — he simply replaced his old annotations on the right with new ones that now embodied his new quantum theory of scattering. I think this is a wonderful historical example of what every physicist knows: Theories come and go, but good experimental data never dies.

![Fig. 5. Compton’s vector diagram for the interaction between a light quantum and an electron in accordance with his quantum theory of scattering. Source: Compton, “Quantum Theory” (Ref. [25]), p. 486; 385.](image)

We see, in sum, that Compton’s discovery was the culmination of six years of experimental and theoretical research, from 1916–1922. His thought evolved along with his own experimental and theoretical work, in a largely autonomous fashion. There is no indication, in particular, that Compton ever read Einstein’s light-quantum paper of 1905: He did not cite Einstein’s
paper in his own paper of 1923, nor did he even mention Einstein’s name in his own paper. This is in striking contrast to Peter Debye, who proposed the identical quantum theory of scattering virtually simultaneously and independently [26], and who explicitly stated that his point of departure was Einstein’s concept of “needle radiation”.

The chronology of Compton’s and Debye’s work, in fact, is instructive and was as follows:

*November 1922*: Compton reported his discovery to his physics class at Washington University.

*December 1 or 2, 1922*: Compton reported his discovery at a meeting of the American Physical Society in Chicago.

*December 10, 1922*: Compton submitted his paper on his quantum theory of scattering to the Physical Review.

April 15, 1923: Debye’s paper was published in the Physikalische Zeitschrift.

May 1923: Compton’s paper was published in the Physical Review.

Thus, although Compton submitted his paper for publication three months before Debye submitted his paper for publication, Compton’s paper actually appeared in print one month after Debye’s. This led quite a number of physicists at the time, particularly in Europe, to refer to the discovery as the Debye–Compton effect — a practice that Arnold Sommerfeld, who happened to be a visiting professor at the University of Wisconsin at the time and who was aware of Compton’s priority in both experiment and theory, was especially instrumental in quelling. Much later, in an interview in 1962 [27], Debye himself insisted that it should be called the Compton effect, maintaining that the person who did most of the work should get the name.

The above chronology is significant for another reason, namely, that it should be supplemented with an additional entry, as follows:

December 6, 1922: Compton submitted his paper on the total internal reflection of X-rays from glass and silver mirrors to the Philosophical Magazine [28].

Now, there is nothing more characteristic of electromagnetic waves than total internal reflection, and there is nothing more characteristic of light quanta than the Compton effect. Thus, within the space of a single week, Compton reported conclusive experimental evidence for both the wave and the particle nature of X-rays. I regard this as symbolic of the profound dilemma that physicists faced at this time over the nature of radiation.

Compton’s experimental results, however, did not go unchallenged. In October 1923, George L. Clark, working in William Duane’s laboratory at Harvard University, announced that he could not find evidence for the change in wavelength of X-rays that Compton had reported [29]. This was a serious experimental challenge to Compton’s work; it prompted two public debates between Compton and Duane at professional meetings and an exchange of visits to their laboratories. It was not resolved for over a year, until December 1924, when Duane conceded at a meeting of the American Physical Society that Clark’s and his experimental results were faulty.

That resolved the experimental question, but the theoretical question still remained open. Thus, in early 1924 Niels Bohr and Hendrik A. Kramers embraced John C. Slater’s concept of virtual radiation, but excluded Slater entirely from the writing up of the resulting Bohr–Kramers–Slater theory [30]. Slater had accepted the physical reality of Einstein’s light quanta, but
Bohr never had. Most recently, Bohr had declared in his Nobel Lecture in December 1922 that:

In spite of its heuristic value, . . . the hypothesis of light quanta, which is quite irreconcilable with so-called interference phenomena, is not able to throw light on the nature of radiation [31].

The essential consequence of the Bohr–Kramers–Slater theory for the Compton effect was that it predicted that energy and momentum are conserved only statistically in the interaction between a light quantum and an electron. Fortunately, as Charles D. Ellis remarked, “it must be held greatly to the credit of this theory that it was sufficiently precise in its statements to be disproved definitely by experiment” [32]. As indeed it was: A year after it was proposed, on April 18 and 25, 1925, Walther Bothe and Hans Geiger in Berlin reported experiments in which they found coincidences between the scattered light quantum and recoil electron, thus proving that energy and momentum were conserved in the interaction between them [33]. Then, two months later, in June 1925, Compton (now in Chicago) and his student Alfred W. Simon reported even more conclusive coincidence experiments that supported his quantum theory of scattering [34]. Earlier, on April 21, 1925, just after Bohr learned about the Bothe–Geiger results, he appended a postscript to a letter to Ralph H. Fowler in Cambridge saying, “It seems therefore, that there is nothing else to do than to give our revolutionary efforts as honourable a funeral as possible” [35].

Of course, as Einstein wrote in a letter of August 18, 1925, to his friend Paul Ehrenfest: “We both had no doubts about it” [36].

REFERENCES


