Grey Alien Transcript: *“Quantum Theory”* by David Bohm, Chapter 2;
The Photoelectric Effect
AnthonyMakesVideos

1 Introduction

Bonjour tout le monde. Today’s video will continue the coverage of David Bohm’s 1951 book *“Quantum Theory”* starting at the beginning of Chapter 2 discussing the Photoelectric effect [1, Ch. 2, pp. 23–33].

Towards the beginning of the previous episode it was briefly mentioned that Christiaan Huygens inaugurated the wave theory of light [2],
Isaac Newton had then supplanted this with a corpuscular particle theory of light [3],
yet by the beginning of the twentieth century when figures such as Max Planck were on the scene, the wave theory of light was again in vogue.

Before Maxwell, Augustin-Jean Fresnel published in early 1822 a work titled *“De la Lumière”* [4] which covered a wave based theory of the nature of light, and focused on the optical phenomena of diffraction, polarization, and birefringence.

This work was received at l’Académie des Sciences by figures including
André-Marie Ampère, Joseph Fourier, Siméon Denis Poisson,
and Pierre-Simon Laplace [5, p. 261n, p. 369n].

Decades later, James Clerk Maxwell had subsequently made the critical prediction that electromagnetic waves and light have the same speed [6].

Maxwell’s prediction was experimentally vindicated [7], which convincingly strengthened the scientific consensus that light is an electromagnetic wave phenomenon.

Thus at the end of the nineteenth century, the wave theory of light had strong support for its wide acceptance, but this paradigm would soon be challenged in the beginning of the twentieth century.

As for some of the earliest origins of quantum theory,
in 1900 Max Planck advanced his theory on the quantization of radiation oscillators
whose associated Planck distribution came to the forefront for its mathematical fit with data on black-body radiation [8].

Planck's distribution, as discussed previously, had avoided the ultraviolet catastrophe evident in the Rayleigh-Jeans law, and through hypothesizing the quantized energy of radiation oscillators in the black-body walls alluded to an explanatory mechanism for its success.

The then-contemporary approximation of Wien provided a reasonably close fit to the black-body data that also avoided the ultraviolet catastrophe [9], but did not offer the quantum hypothesis that Planck's work included.

As mentioned in the previous episode, Einstein in his 1906 paper *"Planck's theory of radiation and the theory of specific heat"* [10] would proceed to apply Planck's quantum hypothesis to the failure of the classical Dulong-Petit law and the equipartition theorem when considering the specific heat capacity of substances approaching cryogenic temperatures.

The Einstein-solid model would largely succeed in applying Planck's quantum hypothesis to an outstanding problem in classical physics;
although this was in spite of its discrepancies at very low temperatures on the order of 10° K;
discrepancies which would be rectified by the 1912 Debye model [11] which introduced the phonon quasi-particle.

The arc of these advances seems to illuminate a narrative of the early development of quantum theory starting from black-body radiation with Planck up through the specific heat of substances with Einstein, but as will be discussed in this video, this narrative is still incomplete.

For instance, shortly after making his quantum hypothesis of black-body radiation, Planck himself admitted that the radiation oscillators he introduced were theoretical constructs that did not necessarily need to exist so long as their properties were consistent with thermodynamic and electrodynamic results [12, p. 135].

Indeed, in spite of the later application by Einstein of Planck's hypothesis to investigating the theory of specific heat, ultimately Planck's concept of quanta was originally an *ad hoc* mathematical hypothesis primarily intended to match experiments.

Note that this was limited to experiments concerning theoretical black-body radiation, since experiments dedicated to directly measuring Planck's constant from the relation $E = nhv$ had not been carried out at the time of Planck's initial publication.

Planck's constant, now regarded as a universal constant of nature, was only hesitatingly adopted following its nascent presentation as a heuristic explanation [13, p. 159].

More support for what is now established quantum theory would come about following Einstein's four *Annus Mirabilis* papers published in 1905, specifically the first of these on the photoelectric effect.

To recap, these four papers in the order of their publication in 1905 were
"Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt" or
"On a Heuristic Point of View about the Creation and Conversion of Light" on the photoelectric effect [14, 15],
"Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen" or
"Investigations on the theory of Brownian Movement" on Brownian motion [16],
"Zur Elektrodynamik bewegter Körper" or
"On the Electrodynamics of Moving Bodies" on special relativity [17],
and *"Ist die Trägheit eines Körpers von seinem Energeinhalt abhängig?"* or
"Does the Inertia of a Body Depend Upon Its Energy Content?" on mass-energy equivalence [18].

Today, three of the four of these *Annus Mirabilis* papers will be discussed in addition to the scheduled contents from Bohm's *"Quantum Theory"*, namely Einstein's papers on the photoelectric effect, special relativity, and mass-energy equivalence.

As a side note, the paper not being discussed today on Brownian motion is arguably the most used of the four in direct applications today [1].

2 Historical Precedent for Einstein's Photoelectric Hypothesis

The experimental data and results setting the precedent for Einstein's photoelectric effect were accumulating towards the end of the 1800s largely due to work involving cathode rays and the emerging notion of the electron particle in the model of the atom.

Heinrich Hertz, who famously validated Maxwell's theory of electromagnetism throughout his experimental work, made one of the first observations of the photoelectric effect during his experiments on the production and reception of electromagnetic waves [19].

Hertz' apparatus consisted of disjoint spark gaps each in the transmitter and the receiver respectively, whereby the successful transmission of an electromagnetic wave across space would be demonstrated by the appearance of a spark in the receiver.

During the course of troubleshooting the apparent spark detection, Hertz would place his receiver variously in a darkened box or would place panels of various materials including glass and quartz between the transmitter and receiver.

Hertz noticed that ultraviolet light would diminish the size of the resulting spark in the receiver across various combinations of whether the receiver was in the box or not or whether the imposed panel was of a material that would absorb UV radiation or not.

Although this is now recognized as a result of the photoelectric effect, Hertz did not follow up on these observations.

---

[1] As of May 18, 2026 per Google Scholar, the Photoelectric paper has 73 citations, the Brownian Motion paper has 9953 citations, the Special Relativity paper has 9339 citations, and the Mass-Energy Equivalence paper has 759 citations.

The previous episode mentioned that Aristotle once famously quipped that "Nature abhors a vacuum" [2], yet in 1650 Otto von Guericke built the first vacuum pump and demonstrated the first vacuum[3] which marked the beginning of modern thermodynamics.

The invention of the vacuum pump thus enabled experiments involving passing high voltage electricity through low pressure air or vacuums.

Michael Faraday in 1838, for instance, noticed that a light arc would occur between two metal electrodes when a high voltage was applied in a glass tube that was partially evacuated of air [23].

In the late 1850s and early 1860s, Julius Plücker, Johann Hittorf, and Heinrich Geissler would oversee the improvement of the pressure reduction in the glass tubes involved in their experiments, obtaining a pressure of around 10^{-3} atmospheres, and thereby a glow discharge filling the tube rather than merely an arc would occur [24].

From around the late 1860s to 1875, William Crookes would further improve the partial vacuum in these glass discharge tubes to obtain a pressure below 10^{-6} atmospheres [25, 26], and this would enable the early investigation into the notion of electrons as particles and the phenomenon of the photoelectric effect.

A dark space just in front of the cathode of these gas discharge tubes would form, and it was noticed that as more air was pumped from the tube, the dark space would spread out.

These dark spaces are called "Faraday dark spaces" or "Crookes dark spaces".

It is now understood that electrons would leave the cathode and collide with gas atoms within the glass tubes, thus resulting in their characteristic glow.

---

[2] Aristotle never actually wrote the phrase "nature abhors a vacuum". Instead, the phrase is a popular modern distillation of his philosophy and foundational views on physics, formally translated into Latin as *natura abhorret vacuum* by the 16th-century French writer François Rabelais [20].

[3] von Guericke's 1650 vacuum pump was first reported by Gaspar Schott in 1657 [21], and subsequently by von Geuricke himself in 1672 [22].

As air is pumped out of the tube, electrons could travel farther on average before such a collision with a gas atom would occur, until the time the tube is mostly dark, whereby electrons could travel in straight lines from the cathode to the anode without a collision.

Therefore, cathode rays consist of electrons accelerated to high velocities traveling from the cathode to the anode and are not apparent in the human visible light spectrum.

This conception of cathode rays as electrons was not immediately understood towards the end of the nineteenth century.

In addition to creating cathode rays directly from an induced voltage, it was also acknowledged around this time from cathode ray experiments that shining ultraviolet light on the cathode of a gas discharge tube also produced a current with the same properties associated with cathode rays.

J.J. Thomson and his contemporary Philipp Lenard, who in his own right was recognized as an expert on cathode rays, were at the forefront of exploring this phenomenon.

The understanding of cathode rays at the time was uncertain whether the emission of negative electricity was a fluid, a molecule, an ion, or waves in the ether, and in many ways impinged upon gaining a better understanding of the nature of electricity.

For Thomson and Lenard, this amounted to determining the charge-to-mass ratio of the contents of cathode rays due to an applied voltage source and cathode rays created by ultraviolet light.

From 1897 to 1898, Thomson would identify that cathode rays are composed of discrete particles and measured the mass of these cathode ray constituents, which he referred to as “corpuscules” but were subsequently known as “electrons” [27].

2.1 Phillip Lenard’s Cathode Ray Experiments

In 1899 Lenard would carry out related experiments seeking to determine the charge-to-mass ratio of the carriers of electricity resulting from cathode rays.

However, when Lenard went to publish the paper [28] containing his results he found that Thomson had already published his own work on cathode rays [27] a few weeks earlier.

Lenard continued by focusing further experiments on how the energy of the emitted photoelectrons varied with the intensity of incident light [29].

The purpose of Lenard's 1902 work was not as much focused on examining the photoelectric effect but rather investigating the distribution of electron velocities within the atom.

Lenard made three key observations through these experiments, namely:

1. The maximum velocity of the ejected electrons depended on the type (or the "spectral composition") of the light employed.
2. The maximum velocity of the electrons departing from a metal plate when illuminated with ultraviolet light remains unaffected by changes in light intensity.
3. There exists a maximum velocity for any given cathode material.

The surprising result that the independence of the maximum velocity of the ejected electrons with respect to the intensity of the incident ultraviolet light did not cause Lenard to speculate about the nature of light,
but was rather the basis for him to attempt to interpret the structure of the atom.

The critical data from Lenard's 1902 paper depicts on the horizontal axis the potential difference between the metal plate receiving the ultraviolet light and the metal plate that the ejected electrons would accelerate towards,
and on the vertical axis the charge per unit time normalized to the saturation current,
with different traces on the plot corresponding to using different materials as photocathodes.

The current emitted by the surface was determined by the intensity of the applied light, i.e., doubling the intensity of the light doubled the number of electrons emitted from the surface.

(I.e., as long as the light was already above the frequency threshold that would allow electrons to be emitted from the surface in the first place.)

That the energy of the emitted electrons was independent of the applied light intensity appeared to contradict Maxwell's prevailing wave theory of light, which predicted that electron energy would instead be proportional to the intensity of the incident light.

2.2 Einstein's Photoelectric Effect Paper

The story from here so goes that the first of Einstein's *Annus Mirabilis* papers in 1905 would aim to explain Lenard's experimental data from observations of the photoelectric effect by advancing a hypothesis inspired by Planck's quantum hypothesis wherein Einstein theorized that light energy is carried in discrete quantized packets.

Einstein's so-called photoelectric effect paper begins in its introduction section by saying

... [According] to Maxwell's theory, the energy of purely electromagnetic phenomena (such as light) should be represented by a continuous function of space.

By contrast, the energy of a material body should be represented by a discrete sum over the atoms and electrons; hence, the energy of a material body cannot be divided into arbitrarily many, arbitrarily small components.

However, according to Maxwell's theory (or, indeed, any wave theory), the energy of a light wave emitted from a point source is distributed continuously over an ever larger volume.

The wave theory of light with its continuous spatial functions has proven to be an excellent model of purely optical phenomena and presumably will never be replaced by another theory.

Nevertheless, we should consider that optical experiments observe only time-averaged values, rather than instantaneous values.

Hence, despite the perfect agreement of Maxwell's theory with experiment, the use of continuous spatial functions to describe light may lead to contradictions with experiments, especially when applied to the generation and transformation of light.

In particular, black-body radiation, photoluminescence, generation of cathode rays from ultraviolet light and other phenomena associated with the generation and transformation of light seem better modeled by assuming that the energy of light is distributed discontinuously over space.

According to this picture, the energy of a light wave emitted from a point source is ***not*** *spread continuously over ever larger volumes, but consists of a finite number of energy quanta that are spatially localized at points of space, move without dividing and are absorbed or generated only as a whole* [15].

A closer reading of the contents of the paper, however, suggests that it primarily deals with how the volume dependence of the entropy of radiation in Planck's black-body law is the same as that of spatially localized independent particles.

Einstein's photoelectric paper is actually quite similar to the core of the first chapter of Bohm's *"Quantum Theory"* [1, Ch. 1, pp. 5–21], as it begins by sketching an expression for black-body equilibrium energy similar to the Rayleigh-Jeans law and noting that in the limit of greater radiation frequency, this expression diverges.

The first main section of Einstein's photoelectric paper says

Let there be a cavity with perfectly reflecting walls, filled with a number of freely moving electrons and gas molecules that interact via conservative forces whenever they come close, i.e., those collide with each other just as gas molecules in the kinetic theory of gases.

In addition, let there be a number of electrons bound to spatially well-separated points by restoring forces that increase linearly with separation.

These electrons also interact with the free molecules and electrons by conservative potentials when they approach very closely.

We denote these electrons, which are bound at points of space, as "resonators", since they absorb and emit the electromagnetic waves of a particular period.

According to the present theory of the generation of light,m the radiation in the cavity must be identical to black-body radiation (which may be found by assuming Maxwell's theory and dynamic equilibrium), at least if one assumes that resonators exist for every frequency under consideration [15].

This is essentially an equivalent ansatz to what was covered roughly across sections 2, 4, and 9 of chapter 1 of Bohm's *"Quantum Theory"*.

The equivalence is more apparent when one considers that Einstein's paper uses a style of notation emphasizing different physical quantities, notably omitting Planck's constant as will be discussed shortly.

For example, Einstein notates the average energy of a "resonator" in the black-body cavity at equilibrium

$$\bar{E} = \frac{R}{N}T$$

where R is the ideal gas constant, N is the number of "real molecules" in a gram equivalent, and T is the absolute temperature.

Practical use of this relation is possible using the ideal constant 8.3145 J/(mol·K) and then computing N using the molar mass of the gas in question (for Einstein's purposes this is probably atomic hydrogen, which would be approximately 1.008 g/mol) and using Avogadro's number 6.022×10^{23} molecules per mol.

2.3 Einstein's Photoelectric Paper vs. Chapter 1 of Bohm's *"Quantum Theory"*

Bohm denotes the same quantity resulting from the equipartition theorem by

$$\bar{E} = \kappa T$$

where κ is Boltzmann's constant, namely 1.3806×10^{-23} J/K.

Bohm's demonstration of the inadequacy of the Rayleigh-Jeans law in a textbook form is more detailed in terms of its derivation from Maxwell's laws than the sketch provided in Einstein's photoelectric paper, but both compute equivalent statements for the energy density of the cavity radiation for frequencies between v and $v + \mathrm{d}v$.

Einstein denotes this quantity in his paper $\rho_v \mathrm{d}v$, and defines

$$\rho_v = \frac{R}{N}\frac{8\pi v^2}{L^3}T$$

where L in this case is the speed of light.

Einstein then explicitly writes the integral statement for the divergence of the black-body radiation according to this relation

$$\int_0^\infty \rho_v \ \mathrm{d}v = \frac{R}{N}\frac{8\pi}{L^3}T\int_0^\infty v^2 \ \mathrm{d}v = \infty.$$

Bohm conversely computes the total number of oscillators between v and $v + \mathrm{d}v$ given by

$$\delta N = \frac{8\pi V}{c^3}v^2 \ \mathrm{d}v$$

denoting the speed of light with the now standard convention as c, and V is the volume of the black-body cavity which in Bohm's case is assumed to be cubic.

Bohm shortly thereafter collects his results and makes his equivalent statement for the Rayleigh-Jeans law

$$U(v) \ \mathrm{d}v = \bar{E} \ \delta N = \frac{8\pi V}{c^3}\kappa T v^2 \ \mathrm{d}v$$

and Bohm does not explicitly write an integral expressing the divergence of this distribution as Einstein did.

Implicit in this comparison is Bohm's hand-wave away of V along the lines of asserting that his theoretical black-body cavity is of unit volume, and the conversion of Bohm's use of the Boltzmann constant to terms of the ideal gas constant which is given by $\kappa = R/N_A$.

Einstein meanders for another 5 sections and then at the end of the section titled *"Interpretation of the Volume Dependence of the Entropy of Monochromatic Radiation using Boltzmann's Principle"*, he says the following:

Subsequently we conclude: In terms of heat theory monochromatic radiation of low density (within the realm of validity of Wien's radiation formula) behaves as if it consisted of independent energy quanta of the magnitude $R\beta v/N$.

We also want to compare the average magnitude of the energy quanta of the "black-body radiation" with the mean average energy of the center-of-mass-motion of a molecule at the same temperature.

The latter is $\frac{3}{2}(R/N)T$, *and for the average energy of the Energy quanta Wien's formula gives:*

$$\frac{\int_0^\infty \alpha v^3 e^{-\frac{\beta v}{T}}\,\mathrm{d}v}{\int_0^\infty \frac{N}{R\beta v}\alpha v^3 e^{-\frac{\beta v}{T}}\,\mathrm{d}v} = 3\frac{R}{N}T.$$

The fact that monochromatic radiation (of sufficiently low density) behaves as regards to dependency of entropy on volume like a discontinuous medium that consists of energy quanta of magnitude $R\beta v/N$ *suggests we should investigate whether the laws of generation and transformation of light are what they must be if light consisted of such energy quanta* [15].

In the above, $\alpha = 6.1 \times 10^{-56}$ and $\beta = 4.866 \times 10^{-11}$.

Both of these constants were derived from Planck's formula as given in Planck's 1901 work titled *"On the Law of Distribution of Energy in the Normal Spectrum"* [30].

More on the Planck constant will be discussed following the current overview of the contents of Einstein's photoelectric paper.

A similar calculation was made by Bohm, but his calculation was at the climax of his argument demonstrating the inadequacy of the classically justified Rayleigh-Jeans law.

Toward the end of section 10 of chapter 1 of *"Quantum Theory"*, Bohm had a statement for the normalized probability that the energy of each black-body radiation oscillator lies between E and $E + \mathrm{d}E$:

$$W(E)\ \mathrm{d}E = \frac{e^{-E/(\kappa T)}\ \mathrm{d}E}{\int_0^\infty e^{-E/(\kappa T)}\ \mathrm{d}E}$$

This was then used to obtain the mean value of the energy $\bar{E}$ shown earlier by integration of $EW(E)$ over all energies to get

$$\bar{E} = \frac{\int_0^\infty Ee^{-E/(\kappa T)}\ \mathrm{d}E}{\int_0^\infty e^{-E/(\kappa T)}\ \mathrm{d}E} = \kappa T$$

thus showing that the average energy of each oscillator is κT.

As just mentioned, the Boltzmann constant can be interchanged with the ideal gas constant by invoking $\kappa = R/N_A$, and thus Bohm's statement and Einstein's statement are seen to be equivalent.

Nonetheless, Einstein's result comes at the end of the sixth substantive section of his paper, and in light of the content of the previous video covering chapter 1 of Bohm's *"Quantum Theory"*, it is hopefully apparent how the paper mostly deals with results related to Planck's black-body radiation law.

2.4 Einstein's Photoelectric Emission Hypothesis

Of the 3 remaining sections of Einstein's photoelectric paper, the sections titled *"Stokes' Rule"* and *"On the Generation of Cathode Rays by Illumination of Solid Bodies"* contain the essence of Einstein's photoelectric hypothesis.

Note that what is referred to in Einstein's paper as "Stokes' Rule" here is also known as Stokes shift which is the difference between positions (expressed in terms of energy, wavenumber, or frequency) of the spectral band maxima of the absorption and emission spectra of the same electronic transition [31].

Starting in the "Stokes' Rule" section, Einstein writes:

Let monochromatic light be transformed by photoluminescence into light of another frequency, and let it be assumed that according to the result just obtained the generating as well as the generated light consists of energy quanta of magnitude $(R/N)\beta v$*, where* v *is the corresponding frequency.*

The transformation process can then be interpreted as follows.

Each generating energy quantum of frequency v_1 *is absorbed and generates–at least with sufficiently small density of the generating energy quanta–by itself a light quantum of frequency* v_2*; possibly other light quanta of frequency* v_3*,* v_4*, etc. as well as other forms of energy (e.g., heat) can be generated simultaneously.*

Through which intermedia processes the final result comes about is immaterial.

If the photoluminescing substance isn't a continuous source of energy it follows from the energy principle that the energy of the generated energy quanta are not larger than the generating light quanta; therefore the following relation must hold:

$$\frac{R}{N}\beta v_2 \leq \frac{R}{N}\beta v_1$$

or

$$v_2 \leq v_1.$$

As is well known this is Stokes' rule [15].

In the next section *"On the Generation of Cathode Rays by Illumination of Solid Bodies"*, Einstein continues:

The usual understanding, that the energy of light is distributed over the space through which it travels in a continuous way encounters extraordinarily large difficulties in attempts to explain photoelectric phenomena, as has been presented in the groundbreaking article by Mr. Lenard.

According to the understanding that the exciting light consists of energy quanta of energy $(R/N)\beta v$ the generation of cathode rays by light can be conceived as follows.

Quanta of energy penetrate the surface layer of the solid, and their energy is transformed, at least partially, in kinetic energy of electrons.

The simplest picture is one where the light quantum gives its entire energy to a single electron; we assume that this will occur [15].

Einstein then says:

[We] must assume that on leaving the solid every electron must do an amount of work P (characteristic of that solid).

Electrons residing right at the surface, excited at right angles to it, will leave the solid with the largest normal velocity.

The kinetic energy of such electrons is

$$\frac{R}{N}\beta v - P. \quad [15]$$

Further down in the section, Einstein then makes what is arguably his fundamental observation concerning the photoelectric effect.

If what he has said thus far is correct, then the maximum kinetic energy per unit charge,

as a function of frequency of the excited light represented in Cartesian coordinates, must be a straight line, whose inclination is independent from the nature of the substance investigated [15].

2.4.1 Einstein’s photoelectric emission hypothesis as an interpretation of Lenard’s data

Recall that Lenard made three key observations through his experiments with cathode rays and ultraviolet light, namely:

1. The maximum velocity of the ejected electrons depended on the type (or the “spectral composition”) of the light employed.
2. The maximum velocity of the electrons departing from a metal plate when illuminated with ultraviolet light remains unaffected by changes in light intensity.
3. There exists a maximum velocity for any given cathode material.

Einstein’s photoelectric hypothesis explains all three of Lenard’s experimental observations.

First, consider the observation that the maximum velocity of the ejected electrons depended on the type (or the “spectral composition”) of the light employed.

Spectral composition alludes to Einstein’s mention of monochromatic radiation, which is radiation of only one particular wavelength or frequency.

By using monochromatic radiation, the effects of radiation frequency can thereby be isolated.

In the case of the maximum velocity of the ejected electrons being zero, then this is equivalent to $\frac{R}{N}\beta v$ (or in modern terms hv) being less than the work function of the metal, which Einstein denotes P but Bohm denotes W.

Only above a threshold frequency will electrons be ejected.

The energy of the ejected electrons is given by $\frac{1}{2}mv^2$, and according to Einstein’s photoelectric hypothesis, this energy is equal to a linear function of the frequency of the incident light.

Second, consider the observation that the maximum velocity of the electrons departing from a metal plate when illuminated with ultraviolet light remains unaffected by changes in light intensity.

Light intensity in strict terms is the power transferred per unit area of a plane perpendicular to the direction of energy propagation.

If the energy of the ejected electron only depends on the transfer of energy to it from a quanta of light, then the energy of the electron is independent of the light intensity.

If the incident light is below the threshold frequency needed to overcome the work function of the material, then increasing the light intensity merely amounts to increasing the number of photons with an inadequate frequency to eject an electron that are bombarding the material.

This explains why after light below the photoelectric threshold is applied for a long time, even at high intensity, electron emission doesn't occur after energy would classically be expected to "build up", and similarly why light above the threshold photoelectric frequency at low intensity immediately causes electron emission.

And third, that there exists a maximum velocity for any given cathode material is due to its characteristic work function which Einstein accounted for in a statement from his photoelectric paper that can be denoted more modernly as $K_{max} = hv - W$.

2.4.2 Robert Millikan's verification of/attempt to disprove Einstein's photoelectric hypothesis

After Einstein's paper was published in 1905, Robert Millikan was convinced that Einstein was wrong and spent the next decade attempting to rigorously test Einstein's photoelectric hypothesis.

Millikan summarized the assertions of Einstein's photoelectric hypothesis as:

1. There is a linear relation between the frequency of the impressed light and the maximum energy of the emission of the electrons ejected by it.
2. The slope of the line representing the linear relation between the potential difference or stopping voltage and the velocity of the ejected electron is h/e i.e., this slope times the charge of the electron e is Planck's constant h.
3. The intercept of the potential difference line on the velocity axis gives the frequency v_0 at which the metal in question first begins to be photoelectrically active.

In 1909, Millikan had conducted the oil drop experiment to measure the charge of a single electron [32, 33], and by 1914 confirmed all three of the assertions of Einstein's photoelectric hypothesis in his paper titled *"A Direct Determination of h"* [34].

With the experimental measure of a single electron in hand, this also meant that Millikan's 1914 work included a measurement of Planck's constant from his photoelectric emission data.

Two years later in 1916, Millikan would publish a follow up work titled *"A Direct Photoelectric Determination of Planck's h"* [35].

In spite of opining in the beginning of this paper that Einstein's original 1905 photoelectric hypothesis was reckless,
Millikan's work experimentally verified that the hypothesis was nonetheless rigorously correct
and in the 1916 paper Millikan again provided a measurement of Planck's constant.

It can perhaps thus be understood that because of the lag between Planck's original publication of his quantum hypothesis of black-body radiation in 1900 and the publication of the experimental measurement of Planck's constant h in 1914,
this explains why Einstein's original photoelectric paper does not use Planck's constant while later treatments of strikingly similar material readily uses Planck's constant
such as has been discussed so far from Bohm's *"Quantum Theory"*.

In 1921, Einstein was awarded the Nobel Prize in Physics, cited "for his services to Theoretical Physics, and especially for his discovery of the law of the photoelectric effect" [36].

Two years later in 1923, Millikan was awarded the Nobel Prize in Physics, cited "for his work on the elementary charge of electricity and on the photoelectric effect" [37].

3 Tenets of the Copenhagen Interpretation embodied by Black-body Radiation and the Photoelectric Effect

At this point, 3 out of the 5 core tenets of the so-called Copenhagen Interpretation of Quantum Mechanics á la Bohr and Heisenberg are beginning to become apparent.

The two tenets not yet covered at this point in the second chapter of Bohm's *"Quantum Theory"* are superposition and the collapse of the wave function.

The three tenets of the Copenhagen Interpretation that have emerged through the discussion of Planck's quantum hypothesis of black-body radiation and Einstein's photon hypothesis of the photoelectric effect are the correspondence principle, the probabilistic nature of the quantum laws, and complementarity.

The correspondence principle was first set forth by Bohr stating that the laws of quantum theory must be so chosen that in the classical limit, where many quanta are involved, the quantum laws lead to the classical laws as an average or in the limit.

This tenet is a reconciliation of the uniquely quantum phenomena observed at scales vanishingly smaller than the macroscopic scale at which the classical laws prevail.

The probabilistic aspect of quantum processes is at odds with the completely deterministic and continuous nature of classical theory in physics.

For example, the classical description of the trajectory of an electron is determined by its position at any instant of time, the velocity at that time, and the value of the force vector acting upon that electron at all times, which force is entirely due to electric and magnetic fields which can be calculated exactly with Maxwell's equations and the initial values of the electric and magnetic fields at all locations.

The operation of this complete determinism is predicated upon the metaphysical framework of gradual and continuous processes in classical theory.

Whereas classical theory completely determines the relationship between variables at an earlier time and those at a later time (i.e., it is completely causal), quantum laws determine only the probabilities of future events in terms of given conditions in the past.

The overlap of the correspondence principle and the probabilistic nature of quantum laws can be understood by analogy to the lifetime of a person.

Population statistics, for example as used for insurance,
can predict the *mean* lifetime of an individual person within a large group
or provide upper bounds such as that it is highly unlikely that an individual will live longer than 120 years,
but it is nevertheless impossible to *exactly predict* the lifetime of the same individual in the group using the same statistics.

Complementarity is in some respects the result of the scope of human observation being limited by default to a macroscopic scale, and thus it is necessary for humans to make measurements or inferences involving apparatuses configured so that quantum states in the system of interest are correlated with a disturbance at the classical macroscopic level of an observable system
in order to glean insights towards the quantum state that are subject to measurement uncertainty.

As demonstrated by the results of the photoelectric effect and other concepts and experiments to be discussed later such as the Franck-Hertz experiment, the Compton wavelength, and the Stern-Gerlach experiment,
electrons and light seem not have a stable self-identity,
but rather, through the interactions with measuring devices, these phenomena manifest wave-like or particle-like behavior contingent upon the details of the specific interactions and observations taking place.

Complementarity in these terms is a more subtle concept which encompasses the notion of wave-particle duality.

3.1 Black-body Radiation under the Copenhagen Interpretation

From the previous chapter, Planck's quantum hypothesis of black-body radiation speaks directly to the probabilistic character of quantum laws and the correspondence principle.

Planck's hypothesis of the quantization of the energy was employed to calculate the average energy emitted by radiation oscillators by determining the likelihood that the oscillator occupies a particular n^{th} discrete energy level, with higher energy levels being less probable.

The first step in the derivation provided by Bohm in chapter one was a statistical mechanical argument for choosing this probability, namely the distribution $e^{-E/\kappa T}$, which was then modified to obtain Planck's distribution by incorporating the discrete energy states given by $E = nhv$.

The resulting distribution was thus $\bar{E} = \dfrac{hve^{-hv/\kappa T}}{1 - e^{-hv/\kappa T}}$.

As discussed in the previous episode, when temperature T is high or conversely frequency v is low, $hv \ll \kappa T$ i.e., quantum energy is very small compared to the thermal energy, Planck's law reduces in this limit to $\bar{E} = \kappa T$ which is the classical result in agreement with the Rayleigh-Jeans law being correct for small $hv/\kappa T$.

3.2 The Photoelectric Effect under the Copenhagen Interpretation

The photoelectric effect, on the other hand, is not a macroscopic effect so it is not a corollary to the correspondence principle.

The photoelectric effect is a low quantum number phenomenon and represents an abrupt shift to quantum behavior from high quantum number classical behavior.

Nevertheless, the photoelectric effect still emphasizes complementarity and the probabilistic nature of quantum processes.

The emerging interpretation of the phenomena associated with the photoelectric effect is that light, which above a threshold frequency is capable of ejecting electrons from a material, consists of localized particles of quantized energy that transfer all of their energy to the electrons during a collision.

A conceptual quandary thus emerges between the notion of light as a wave convincingly demonstrated by the success of Maxwell's equations and the notion of light as a photon particle from the foregoing insights due to Einstein.

If the photoelectric effect were interpreted as a classical wave, it might be surmised that when the radiation strikes an electron vibrating within an atom of a material, the radiation transfers energy to the electron.

If the electric field component of the light wave radiation is of a resonant frequency with the frequency of the electron in this atom, the electron would perhaps absorb energy from the light wave until it had gained an amount equal to hv, after which it would be ejected.

This triggering mechanism of the photoelectric effect based on a resonance phenomenon was actually endorsed by Lenard contra Einstein's 1905 photoelectric hypothesis.

Above the minimum frequency at which electron emission occurs, this resonance frequency triggering hypothesis implies that there would be a discrete set of numerous other frequencies at which the emission of electrons would occur.

Rather, this is ruled out in light of the observation that all frequencies above the photoelectric threshold for a material result in the emission of electrons.

Therefore, at the scale of the photoelectric emission of electrons from a material, the photon particle notion of light prevails over the conception of light as waves.

The same cannot be said, however, for explaining the results of the double-slit experiment.

The double-slit experiment and the characteristic interference patterns observed had already been carried out as early as 1801 [38], and these interference patterns cannot be explained by conceiving of light as a photon particle.

Furthermore, the photon explanation of the photoelectric effect does not supplant the macroscopic success of Maxwell's wave-based equations which were already by the twentieth century convincingly accurate for describing the behavior of a whole host of phenomena including radio waves and x-rays.

Apropos the double slit experiment and the photoelectric effect, these results are explained by conceiving of light as either a particle or a wave but not simultaneously both.

An additional difficulty not only apparent in photoelectric effect experiments but others as well is that it is observed that the time and location of the individual quantum transfer cannot be exactly predicted with certainty.

Only the probability of such a process's occurrence may be predicted.

In the case of many quanta, it is possible to predict the average transfer effect in any given region.

In light of its tremendous success quantitatively describing quantum phenomena but its shortcomings with providing metaphysical explanations, the mantra for the Copenhagen Interpretation is "don't think, just calculate".

It is worth mentioning that Bohm's *"Quantum Theory"* is intended to be Bohm's physical motivation for the prevailing Copenhagen Interpretation.

Quantum Theory up to the time when Bohm's book was published in 1951 had been championed by presentations such as that of Paul Dirac's 1930 monograph *"The Principles of Quantum Mechanics"* which is a heavily mathematical treatment [39], for example.

4 Timing of the Photoelectric Effect

4.1 *"The Element of Time in the Photoelectric Effect"* by Lawrence and Beams (1928)

Around the point in chapter two of Bohm's book where the topics discussed thus far have been introduced, Bohm strives to provide physical, experimental justification for considering a quanta of energy as an indivisible unit of energy and the transfer of a quanta from one system to another as an indivisible elementary process.

To do so, Bohm briefly mentions a paper titled *"The Element of Time in the Photoelectric Effect"* by Ernest O. Lawrence and J.W. Beams published in September of 1928 [40], which will now be discussed in slightly more detail than was done in Bohm's book.

The Lawrence and Beams paper describes the experimental setup and methods used by the authors to attempt to study the time variation of photoelectric emission from a metal surface illuminated by light flashes with a duration on the scale of tens of nanoseconds.

These experimental results are then used by Lawrence and Beams to weigh in on the details of the photoelectric emission process by atoms receiving the illumination, thus a property of the nature of atoms,
and additionally used by Bohm to argue that the transfer of a quanta of energy by light to the metal atoms was demonstrated to not have been broken up into smaller quanta and that thus quanta of energy are indivisible units of energy.

To be clear, the 1928 Lawrence and Beams paper is now investigating the photoelectric effect beyond merely the notion that light has a particle-like nature under the circumstances of the photoelectric effect.

The timing precision of Lawrence and Beams' experimental setup is founded upon their through the use of Nicol prisms aligned with Kerr cell shutters.

Nicol Prisms are calcite crystals used to plane polarize ordinary light which is beamed through a configuration of these prisms and the Kerr cell shutters.

Kerr cell shutters consist of a transparent container filled with nitrobenzene attached to electrodes through which a high voltage on the scale of 10-30 kV is passed which produces an electric field perpendicular to the transmitted light beam.

This causes the nitrobenzene container to act like a photographic shutter that can be opened for a brief amount of time on the scale of tens of nanoseconds that can be configured by the known lengths of the wires controlling the electrodes.

This shutter behavior is due to the Kerr effect which is a change in the refractive index of a material in response to an applied electric field.

This effect is also known as the quadratic electro-optic effect because the induced birefringence in the material, which in the case of the Kerr cell is the nitrobenzene, is proportional to the square of the applied electric field.

The light bursts then illuminate the potassium hydride surface of a photoelectric cell, and analog circuits involving spark gaps are used to restrict the current collected from the ejected electrons on a brass ball to be restricted to coincide as closely as possible with the timing of the light bursts.

Correction to the apparent time of cut-off of the electron collection is necessary because of the finite time required for the electrons to pass from the potassium hydride surface to the brass ball collector.

As for the measurement of the collected ejected electrons, a collecting electrode was connected to a Dolezalek electrometer.

A Dolezalek electrometer is a high-sensitivity quadrant electrometer which is a robust variation of a gold-leaf electroscope wherein a charge applied to a pith ball attached to a rod with a gold leaf draped over it in a draft-sealed container.

Bringing a charged object near the pith ball causes the gold leaves to spread apart proportional to the applied charge.

Lawrence and Beams' experiments collected data for the photoelectric currents to their brass ball collector corresponding to various times of the reversal of the electric field between the potassium hydride photoelectric cell surface and the collector after the beginning of the spark discharge timing the initiation of the light burst.

Initial replicates encountered spurious light passing through the shutter before the main flash started.

They subsequently obtained data upon shortening the wire paths to their Kerr cell shutters and varying the bias potential drawing the photoelectrons to troubleshoot cut-off timing variations that they encountered during the course of their experiments.

It is worth noting that after preparing their photoelectrode surface, the various parts became photoelectrically active and so the photoelectric sensitivity of their potassium hydride surface would drift in the course of a few days.

The resulting data was plotted neatly by hand and incorporated into their final publication.

In discussing their data, Lawrence and Beams stated that

There is no appreciable lag in the Kerr effect, but that quite appreciable intervals of time required to discharge the Kerr cells affect the operation of the shutter in such a way that it does not cut off the light sharply at the instant that the second of the Kerr cells begins discharging.

Assuming there exists no persistence of the electron emission after the surface is illuminated, the data indicate that the shutter continued to allow light to pass through for about 10^{-8} *seconds after the second Kerr cell began discharging* [40].

They continued on to say that

> *The uncertainty in the functioning of the shutter makes it impossible to conclude with great precision that the photoelectric emission stops as abruptly with the illumination as it commences.*
>
> *It can be stated with confidence only that the sum of the time required for the double refraction in the Kerr cells to decay to a small value and the time that the photoelectron emission persists after the illumination is cut off is less than* 10^{-8} *seconds* [40].

Furthermore, in their experimental setup, Lawrence and Beams observed that the electron emission began very nearly at the same instant the illumination commenced.

As such, the induced voltage resulted in wave fronts traveling steeply in the range of tens of nanoseconds, and their efforts in this work amounted to being the first definite measurement of the steepness of such wave fronts.

Nevertheless, their setup was plagued by the presence of high frequency oscillations.

These high frequency oscillations were present in their photoelectric cell and other parts of the driving circuit, so

> *[because] of the very small photoelectric emission and the very sudden changes in potential of various parts of the circuit, spurious effects were always in evidence in the early stages of the experimental work and it seemed an impossible task to arrive at trustworthy results* [40].

In spite of these issues, they ultimately asserted that these sources of trouble were eliminated and that the results they presented were entirely reliable.

Their conclusion upon arriving at their final results was that their experiments were unable to detect with certainty a difference between the beginning of the light flash and the beginning of the electron emission less than 3×10^{-9} seconds, and that from the limits of this purview the photoelectric effect is instantaneous.

The results also indicated to the authors that Bohr's contemporary hypothesis of the mechanism of the absorption and emission of light and electrons from atoms was correct.

Namely, when a quanta of light is absorbed an electron is raised to an excited outer orbit of the atom or else if the energy of light exceeds the distinct binding energy of the electrons in the atomic, molecular, or crystalline system the electron is ejected.

Conversely, light is emitted by an atom when excited electrons return to their normal energy level orbitals.

The best experimental data (at the time of the publication of the Lawrence and Beams paper) on atom excitation was mentioned in the paper as having measured atom excitation to take place in less than 10^{-10} seconds [41], and thus combined with Lawrence and Beams' results it is suggested that Bohr's hypothesis is correct.

The paper concludes by saying that

... it is therefore natural to expect that photoelectric emission would persist after illumination of a metal surface.

Because of the uncertainty of the rapidity with which the elctro-optical shutter cut off of the light in the present experiments it is impossible to determine with great precision whether or not such a persistence of the photoelectric effect exists.

It can only be said that such an effect becomes inappreciable within 10^{-8} *seconds after cessation of the illumination* [40].

Therefore, in section three of chapter two of *"Quantum Theory"* when Bohm states that

[Lawrence and Beams] found ... that none of the quanta was ever broken up... [1, Ch. 2, pp. 27]

was not a rigorous inference of their results.

The Lawrence and Beams 1928 paper rather suggests that the process of the transfer of light quanta had not yet been observed on time scales less than tens of nanoseconds,
and therefore if the process of the transfer of light quanta *were* actually divisible then its separation into subprocesses would occur on such a shorter time scale,
so this possibility had not been definitively ruled out by their experimental results.

4.2 *"Absolute Timing of the Photoelectric Effect"* by Ossiander and Riemensberger (2018)

90 years later in September of 2018, German physicists working out of the Technical University of Münich and the Max-Planck Institute for Quantum Optics would publish results [42] that determine the absolute timing sequence of the photoelectric effect in condensed matter systems on the order of literally one billion times more precise than that achieved by Lawrence and Beams.

The experimental method used by first authors Ossiander and Riemensberger measured initial electron packet creation, transport, and scattering on atomic length scales, or Angstroms which are on the order of 10^{-10} meters or one ten-billionth of a meter, and attosecond time scales, which are on the order of 10^{-18} seconds.

The authors used a streak camera arrangement to perform attosecond scale measurements.

In principle, a streak camera uses a cathode-ray tube to accelerate photoelectrons produced from an illuminated photocathode via the photoelectric effect, and then the emitted electrons are deflected by an electric field by a time-varying high voltage between a pair of plates.

The rapidly modulated electric potential between the plates produces a time-varying deflection of the electrons which are subsequently swept across a phosphor screen at the end of the cathode-ray tube.

A linear detector, such as a charge-coupled device or CCD array, is used to measure the streak pattern on the screen representing the temporal profile of the light pulse.

The ultimate aim of the Ossiander and Riemensberger work was to determine the absolute duration of the temporal sequence of events between the arrival time of an ionizing extreme-ultraviolet light at the surface of a solid and photoelectron emission into a vacuum.

To accomplish attosecond precision, adsorbed atoms are used as a chronoscope species whose absolute photoionization timing could be determined in concurrent gas-phase measurements.

Due to a coverage-dependent exit delay for electrons from the particular solid material, data was collected for fractional surface coverage of the chronoscope species in units of saturated monolayers.

Thus, the small influence of this adlayer on the delay over a wide range of coverages was examined so that the behavior of the pristine solid material could be assessed by extrapolation to infer the exit delay as for a bare surface.

For this study, the solid from which photoemission is being stimulated is tungsten 110, which is an isotope containing 74 protons and 110 neutrons, and the adsorbed clock atom species is iodine.

To achieve a true absolute delay timing, gas phase streaking using a mixture of the iodine chronoscope species and helium is undertaken.

Note that helium is inert and thus nonreactive with either tungsten or iodine, so there is no inherent sensitivity drift by the choice of such materials.

When the extreme-ultraviolet light pulse arrives at the surface, electron expulsion is stimulated from the iodine clock atoms, and then the light pulse propagates into the crystal and photoexcites electrons from both localized core 4f orbitals and delocalized conduction bands of the tungsten.

The surface and gauge experiments thus provide absolute timing for the Helium 1s orbital electron emission, photon arrival, the time of exit for an electron from the conduction band of the tungsten, and the time of exit for an electron from the core orbitals of the tungsten.

As the authors of the absolute photoelectric timing paper write:

A major advantage of absolute emission timing is that the duration of the creation process for each observed photoemission feature is individually recorded.

Thus, all delay contributions, even those cancelling in relative measurements, are uncovered by the present absolute measurement, enhancing the importance of the extracted timing information for benchmarking theoretical models and allowing their direct interpretation [42].

The reported delay between photon arrival and the exit of a conduction band electron from tungsten 110 was 40 ± 9 attoseconds.

The delay between photon arrival and the exit of a 4f core orbital electron from the same solid was reported to be 103 ± 6 attoseconds.

These 2018 results timing the photoelectric effect, which as has already been mentioned are on the order of 10^{-18} seconds, are roughly a billion times more precise than the mere upper bound on photoemission of tens of nanoseconds from Lawrence and Beams 90 years earlier as referenced by Bohm in 1951 to physically motivate the indivisibility of the transfer of a quanta of energy as an elementary process.

5 Bohm's "Hidden Variables" Interpretation of Quantum Theory

Switching conceptual gears now, Section 5 of chapter two of Bohm's *"Quantum Theory"* interestingly makes Bohm's first mention in the book of the notion of a "hidden variables" interpretation of quantum theory [1, Ch. 2, pp. 29].

Recall, however, that Bohm's *"Quantum Theory"* book is his description of the Copenhagen interpretation with an emphasis on physical or experimental motivation of the theoretical concepts.

The book *"Quantum Theory"* was published in early 1951 but during the summer of that same year, Bohm submitted a two paper series titled *"A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables"*, parts one [43] and two [44], respectively.

These papers were subsequently published in January of 1952, and they constitute Bohm's assertion of the "hidden variables" interpretation of quantum theory.

Section five of chapter two of Bohm's book, titled *"Unlikelihood of Completely Deterministic Laws on a Deeper Level"* begins by wondering

whether the appearance of probability in quantum processes is not a result of [human] ignorance of the correct variables to use in describing [quantum systems] [1, Ch. 2, pp. 29].

In the book, Bohm then draws the analogy between the probabilistic understanding of quantum theory at the time of his book's writing to thermodynamics.

... [In] thermodynamics... the pressure, temperature, and volume of a given system [are measured].

In very small regions of space... [one finds] that these quantities no longer obey an equation of state exactly, but instead exhibit large random fluctuations about a mean value that is predicted by the equation of state.

Hence, the deterministic laws of thermodynamics break down and are replaced by laws of probability.

This is because the thermodynamic variables are no longer appropriate for the problem and must be replaced by the position and velocity of each molecule, which are, from the viewpoint of thermodynamics, hidden variables.

The thermodynamic quantities are, then, merely averages of hidden variables that cannot be observed by the thermodynamic methods alone.

To find the underlying causal laws, [one] must accept a description in terms of the individual molecules [1, Ch. 2, pp. 29].

Bohm draws the analogy slightly further in section three of the first of his papers by framing the possibility of hidden variables underlying quantum theory in terms of the understanding of early atomic theory.

... The existence of atoms was postulated in order to explain certain large-scale effects, such as the laws of chemical combination, the gas laws, etc.

On the other hand, these same effects could also be described directly in terms of existing macrophysical concepts... and a correct description in these terms did not require any reference to atoms.

Ultimately, however, effects were found which contradicted the predictions obtained by extrapolating certain purely macrophysical theories to the domain of the very small, and which could be understood correctly in terms of the assumption that matter is composed of atoms [43].

Both in the book and his papers, Bohm poses the possibility of hidden variables playing a role in quantum theory:

Perhaps then, our present quantum-mechanical averages are similarly a manifestation of hidden variables, which have not, however, yet been detected directly [43].

However, in the book, Bohm dismisses this possibility as unlikely, saying that

[First,] ... no experiment has yet shown the slightest trace of such hidden variables.

[Second,]... there are strong theoretical arguments which make it unlikely that such hidden variables exist [1, Ch. 2, pp. 29].

Later, at the end of section 19 at the end of chapter 22, the second to last chapter of his book, Bohm continues to say

We conclude then that no theory of mechanically determined hidden variables can lead to **all** *of the results of the quantum theory.*

Such a mechanical theory might conceivably be so ingeniously framed that it would agree with quantum theory for a wide range of predicted experimental results.

> *[Bohm does] not wish to imply here that anyone has ever produced a concrete and successful example of such a theory, but only state that such a theory is, as far as [he knows], conceivable* [1, Ch. 22, pp. 623].

However, in his Hidden Variables papers, Bohm does not make the same conservative rejection of the possibility of an interpretation of quantum theory that is an alternative to the prevailing Copenhagen interpretation.

In the first "Hidden Variables" paper, Bohm writes that

> *If, as is certainly possible, these hidden variables are actually needed for a correct description at small distances, we could easily be kept on the wrong track for a long time by restricting ourselves to the usual interpretation* [i.e., the Copenhagen interpretation] *of the quantum theory, which excludes such hidden variables as a matter of principle* [43].

Bohm reiterates two fundamental assumptions of the Copenhagen interpretation and then recognizes that this interpretation is

> *...tested adequately by the extremely wide range of experiments that are in agreement with predictions obtained by using this system* [43].

However, Bohm continues that these fundamental assumptions of the Copenhagen interpretation do not imply a unique mathematical formulation which could then be experimentally falsified to indicate that these assumptions were wrong.

Rather, the fundamental assumptions of the Copenhagen interpretation merely limit the possible forms of the mathematical theory, and insufficiently so in order to make possible a unique set of predictions that could in principle permit such an falsifiable experimental test.

Therefore, when the Copenhagen interpretation is found to be experimentally inadequate, it is always possible to assume that the theory can be made to agree with experiment by changes in the mathematical formulation alone, not requiring any fundamental changes in the physical interpretation.

What Bohm says next in his *"Hidden Variables"* paper is worth quoting at length.

This means that as long as we accept the usual physical interpretation of the quantum theory, we cannot be led by any conceivable experiment to give up this interpretation, even if it should happen to be wrong.

The usual physical interpretation therefore presents us with a considerable danger of falling into a trap, consisting of a self-closing chain of circular hypotheses, which are in principle unverifiable if true.

The only way of avoiding the possibility of such a trap is to study the consequences of postulates that contradict [the assumptions of the Copenhagen interpretation] at the outset [43].

Therefore, "hidden" elements or variables could be postulated and experiments could be sought that depended in a unique and reproducible way on the assumed state of these hidden elements or variables.

Success would favor the hypothesis that hidden variables exist.

If [said experiments] are not verified, however, the correctness of the usual interpretation of the quantum theory is not necessarily proved, since it may be necessary instead to alter the specific character of the theory that is supposed to describe the behavior of the assumed hidden variables [43].

More can be said about Bohm's "Hidden Variables" interpretation later, but it was perhaps worthwhile to briefly consider some of the philosophical implications of Bohm's thoughts on quantum theory.

6 Derivation of the momentum of a photon

6.1 Bohm's derivation from *"Quantum Theory"*

What has been discussed so far has roughly covered the contents of the first six sections of chapter two of Bohm's *"Quantum Theory"* book.

Section seven of chapter two, titled *"Particle Properties of Light"* considers the particle-like aspects of light in more detail [1, Ch. 2, pp. 31–33].

From a quantum perspective, this particle-like aspect of light follows from the Bohr model of the atom which akin to the previously discussed radiation oscillators can gain or lose energy only by the transfer of a whole quanta at once with the change in energy given by the Planck relation $\Delta E = hv$.

A quantum particle such as a theoretical radiation oscillator excited to the n^{th} quantum state has an energy of $E = nhv$ which can be lost in n steps, where each discrete loss of a quanta of energy equivalently constitutes the emission of a particle with energy hv.

These equivalent particles are *photons*.

From classical electrodynamics, radiation fields possess momentum as well as energy, thus the momentum of these equivalent photon particles can be calculated.

Maxwell's equations show that this momentum is given by

$$\mathbf{p} = \frac{1}{4\pi c} \int (\underline{\mathcal{E}} \times \underline{\mathcal{H}})\ \mathrm{d}\tau.$$

For a light wave in a vacuum, the electric field $\underline{\mathcal{E}}$ is orthogonal to the magnetic field $\underline{\mathcal{H}}$, and the magnitude of the electric field is equal to the magnitude of the magnetic field, i.e., $|\underline{\mathcal{E}}| = |\underline{\mathcal{H}}|$.

Hence, the vector $\underline{\mathcal{E}} \times \underline{\mathcal{H}}$ is orthogonal to both the electric field $\underline{\mathcal{E}}$ and the magnetic field $\underline{\mathcal{H}}$, and therefore is in the direction of the wave propagation vector $\mathbf{k}$.

The magnitude of the cross product of the electric and magnetic fields constituting this light wave is given by

$$|\underline{\mathcal{E}} \times \underline{\mathcal{H}}| = \mathcal{E}^2 = \frac{\mathcal{E}^2 + \mathcal{H}^2}{2}.$$

Since the energy of the wave is given by

$$E = \frac{1}{4\pi} \int \frac{\mathcal{E}^2 + \mathcal{H}^2}{2} \, \mathrm{d}\tau$$

then the momentum becomes

$$\mathbf{p} = \frac{E}{c} \hat{\mathbf{k}} \tag{1}$$

where $\hat{\mathbf{k}}$ is a unit vector in the direction of propagation.

6.2 Relativistic derivation of the momentum of a photon

6.2.1 Special Relativity

Note that a photon travels nominally at the speed of light which is admittedly in a relativistic regime.

As such, this relation for the momentum of a photon can also be obtained from the framework of special relativity.

This derivation relies upon the results of Einstein's third and fourth Annus Mirabilis papers.

These papers are again, namely, *"Zur Elektrodynamik bewegter Körper"* or *"On the Electrodynamics of Moving Bodies"* on special relativity [17], and *"Ist die Trägheit eines Körpers von seinem Energeinhalt abhängig?"* or *"Does the Inertia of a Body Depend Upon Its Energy Content?"* on mass-energy equivalence [18].

The critical metaphysical step made by Einstein in relativity is the abandonment of the absolute frames of reference invoked by Newton, or as Newton calls it "God's boundless sensorium" [3], to describe kinematic motion.

The difference between special and general relativity in short is that general relativity treats observers who can be in accelerated relative motion and in regions of space-time including the action of gravitation.

Thus, special relativity has a restricted range of application for describing motion.

Einstein's special relativity paper was published on June 30, 1905, however it would take him another ten years until he would work out the mathematical details of general relativity which were published in his paper from November of 1915 titled *"Die Feldgleichungen der Gravitation"*, or *"The Field Equations of Gravitation"* [45].

The special relativity paper has two main parts, the first on kinematics and the second on electrodynamics.

To obtain the relativistic derivation of the photon momentum given in Bohm's *"Quantum Theory"*, only the major concepts from the first part on kinematics from Einstein's special relativity paper are needed[4].

6.2.2 Foundations of Special Relativity

There are two postulates at the foundation of special relativity.

The first postulate states that it is impossible to measure or detect the unaccelerated translatory motion of a system through free space or through any ether-like medium which might be assumed to pervade it.

The second postulate states that the velocity of light in free space is the same for all observers, independent of the relative velocity of the source of light and the observer.

These postulates lead to consequences that may seem counterintuitive as to the nature of space and time.

6.2.3 Lorentz Transformation Equations for Position

Consider a source of light S and two systems, A moving *towards* the source S, and B moving *away* from it.

Suppose that observers on the two systems have marked off equal absolute distances aa' and bb' on each of the systems A and B, respectively, in the direction of the source in order to measure the velocity of light by determining the time taken for light to travel from a to a' and from b to b'.

---

[4] Sections 6.2.2 through 6.2.9 follow sections 4 through 14 and 22 through 26 of [46].

Due to the first postulate of relativity, one cannot assign any significance to the absolute velocities of the two systems, but the velocities of the systems relative to the light source S may be considered.

From the second postulate of relativity, the measured velocity of light must be independent of this relative velocity between source and observer.

This leads to the conclusion that the time taken for the light to travel from a to a' shall measure the same as that for the light to travel from b to b', even though the system A is moving towards the source and B is moving away from it.

The resolution for what is seemingly a contradiction lies in the Lorentz transformations, namely the Lorentz contraction in space and time dilation.

Consider two systems of space-time coordinates S and S' in relative motion V, where the systems are provided with sets of rectilinear Cartesian axes, and V is taken for convenience and without loss of generality in the x-direction.

Also, provide each system with a set of clocks distributed at convenient intervals throughout the system and fixed in relative position so as to move with the movement of their respective system.

Let the position in space at which some event occurs be designated by the spatial coordinates x, y, and z with respect to the axes of system S, or the spatial coordinates x', y', and z' with respect to system S'.

Additionally, specify the time at which the event occurs by giving the clock readings t or t' in the respective S or S' systems.

Next, again for convenience and without loss of generality, let the two systems be chosen so that the Cartesian axes OX and $O'X'$ lie along the same line, and for further simplification choose the starting point for time measurements in the two systems so that t and t' are equal to zero when the two origins O and O' are in coincidence.

The problem that now arises is to obtain a set of transformation equations connecting the variables of the two systems that enables the description of any given kinematical occurrence from the variables of the one system to be transformed into those of the other.

The correct transformations were first obtained by Lorentz which he published in a memoir [47], and are hence called the Lorentz transformation equations.

This memoir was unknown to Einstein at the time of the Annus Mirabilis publication of his special relativity paper which provided the same transformations and additionally appreciated their full significance from the point of view of the relativity of motion.

The Lorentz transformations may be written

$$\begin{aligned} x' &= \frac{x - Vt}{\sqrt{1 - V^2/c^2}}, \\ y' &= y, \\ z' &= z, \\ t' &= \frac{t - xV/c^2}{\sqrt{1 - V^2/c^2}} \end{aligned}$$

where V is the relative velocity of the two systems and c is the velocity of light.

Conversely solving for the unprimed quantities in terms of the primed quantities, the Lorentz transformations may be equivalently written in the form

$$\begin{aligned} x &= \frac{x' + Vt'}{\sqrt{1 - V^2/c^2}}, \\ y &= y', \\ z &= z', \\ t &= \frac{t' + x'V/c^2}{\sqrt{1 - V^2/c^2}}. \end{aligned}$$

In accordance with the second postulate of special relativity, the velocity of light must measure the same in both systems of coordinates.

This can be seen by the fact that the following quantity

$$\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 - c^2\mathrm{d}t^2$$

is invariant for the transformation, as shown by substituting the Lorentz transformation equations in terms of unprimed quantities in terms of the primed quantities into this quantity and obtaining an expression of the same form except with primed instead of unprimed quantities.

It is immediately seen from the invariance of this expression that the velocity of light will measure the same in both systems, since if there is an impulse travelling with the velocity c with respect to the system S in accordance with the equation

$$\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 - c^2\mathrm{d}t^2 = 0$$

then one also shall have the light impulse travelling with the velocity c with respect to the system S' in accordance with the equation

$$\mathrm{d}x'^2 + \mathrm{d}y'^2 + \mathrm{d}z'^2 - c^2\mathrm{d}t'^2 = 0.$$

Therefore the Lorentz transformation equations satisfy the two postulates of special relativity.

Note that when the relative velocity between the systems V is small compared with that of light, the Lorentz transformation equations reduce to the so-called Galilean transformation equations

$$\begin{aligned} x' &= x - Vt, \\ y' &= y, \\ z' &= z, \\ t' &= t. \end{aligned}$$

While the Galilean time transformation equation $t' = t$ would indicate a universal time suitable for all observers, the corresponding Lorentz time transformation equation $t' = \dfrac{t - xV/c^2}{\sqrt{1 - V^2/c^2}}$ implies not only that there is no single suitable universal time for all observers, but furthermore that time is not a continuum that is independent of space.

Also note that the set of the Lorentz transformations between all systems in unaccelerated uniform motion constitutes a mathematical group, and therefore the associativity and closure of group elements guarantees that the combined result of successive transformations is equivalent to a single transformation from the original to the final system of coordinates.

Manipulating the Lorentz transformation equations to obtain further equations for transforming the measurements of other geometrical or kinematical quantities depending on the coordinates will permit further physical insights and interpretations.

6.2.4 Differential form of the Lorentz Transformation Equations for neighboring events; Lorentz Contraction and Time Dilation

The foregoing Lorentz transformation equations can be written in terms of the differential quantities dx, dy, dz, dt, dx', dy', dz', and dt':

$$\begin{aligned} \mathrm{d}x' &= \frac{\mathrm{d}x - V\ \mathrm{d}t}{\sqrt{1 - V^2/c^2}}, \\ \mathrm{d}y' &= \mathrm{d}y, \\ \mathrm{d}z' &= \mathrm{d}z, \\ \mathrm{d}t' &= \frac{\mathrm{d}t - \mathrm{d}x\ V/c^2}{\sqrt{1 - V^2/c^2}}. \end{aligned}$$

The differential quantities may be interpreted as giving the measurements in the two systems of the spatial and temporal intervals which correspond to the difference in position and time of some *given* pair of neighboring events.

This enables conclusions to be drawn as to the intercomparison of measuring sticks and clocks in the two systems.

Consider two measuring sticks held parallel to the x-axis, one in each of the two systems, in such a way that their scale divisions can be compared as the two sticks slide past each other;
and consider as the events to be observed the coming into coincidence of division marks on one of the measuring sticks with division marks on the other.

First, determine how a length $\mathrm{d}x'$ laid off on the measuring stick in system S' will appear when measured in the system S.

To do this, consider coincidences, *which appear simultaneous in system* S, between the end points of $\mathrm{d}x'$ and the division marks on the measuring stick in system S.

Since the coincidences are simultaneous in system S, one has that $\mathrm{d}t = 0$, which by substitution into the differential form of the Lorentz transformation equations yields

$$\mathrm{d}x' = \frac{\mathrm{d}x}{\sqrt{1 - V^2/c^2}} \quad \text{or} \quad \mathrm{d}x = \mathrm{d}x'\sqrt{1 - V^2/c^2}.$$

Therefore a measuring stick traveling with system S' and measuring $\mathrm{d}x'$ in the units of that system will measure the shorter length $\mathrm{d}x'\sqrt{1 - V^2/c^2}$ in the units of system S when the simultaneous positions of its ends are observed in that system.

Next, determine how a length $\mathrm{d}x$ laid off on the measuring stick in system S will appear when measured in system S'.

To do this, one must now consider coincidences, *which appear simultaneous in system* S', between the end points of $\mathrm{d}x$ and division marks on the measuring stick in system S'.

Since the coincidences are simultaneous in system S', in accordance with the Lorentz time transformation in differential form as previously given one has

$$\mathrm{d}t' = \frac{\mathrm{d}t - V\ \mathrm{d}x/c^2}{\sqrt{1 - V^2/c^2}} = 0,$$

which when substituted into the x coordinate direction Lorentz transformation equation in differential form yields

$$\mathrm{d}x' = \mathrm{d}x\sqrt{1 - V^2/c^2}.$$

Therefore a measuring stick traveling with system S measures shorter in the same ratio as before when the simultaneous positions of its ends are observed in the other system S'.

In both cases, a measuring stick measures shorter in the ratio of $\sqrt{1-V^2/c^2} : 1$, when moving with the velocity V past the system in which the observation of length is being made, than when measured in a system in which it is at rest.

This result is called Lorentz contraction.

Note that there will be no disagreement as to measurements made in the two systems of coordinates of distances at right angles to the line of motion.

There is no change in length for a measuring stick which is moving perpendicular to its length past the system of coordinates in which it is to be measured.

The differential form of the Lorentz transformation equations can also be used to provide conclusions as to the intercomparison of clocks in relative motion.

First, determine how a time interval $\mathrm{d}t'$ measured on a *single* clock in system S' between two events, which occur at the same point in S', will measure with the clocks of system S.

Since the two events occur at the same point in S', the differential form of the x coordinate equation of the Lorentz transformation yields

$$\mathrm{d}x' = \frac{\mathrm{d}x - V\ \mathrm{d}t}{\sqrt{1-V^2/c^2}} = 0,$$

which upon substitution into the differential form of the time coordinate equation of the Lorentz transformation yields

$$\mathrm{d}t' = \mathrm{d}t\sqrt{1-V^2/c^2} \quad \text{or} \quad \mathrm{d}t = \frac{\mathrm{d}t'}{\sqrt{1-V^2/c^2}}.$$

Therefore, the time duration $\mathrm{d}t'$ when measured with a given clock in system S' will have the longer duration $\mathrm{d}t'/\sqrt{1-V^2/c^2}$ when measured by the clocks in system S.

Next, determine how a time interval $\mathrm{d}t$ which can be measured on a single clock in system S between two events, which occur at the same point in system S, will measure with the clocks of system S'.

For this case where two events occur at the same point in system S one has that

$$\mathrm{d}x = 0$$

which upon substitution into the differential form of the time coordinate equation of the Lorentz transformation yields

$$\mathrm{d}t' = \frac{\mathrm{d}t}{\sqrt{1 - V^2/c^2}}.$$

Again, this shows that the time interval between two events which has the duration $\mathrm{d}t$ when measured with a given clock has a longer duration when measured by clocks relative to which the first clock is moving.

In both cases, the units of time denominated by the single clock appear lengthened in the ratio $1 : \sqrt{1 - V^2/c^2}$ when it is moving with the velocity V past the clocks with which it is being compared.

This result is called time dilation.

6.2.5 The Lorentz Transformation Equations for Velocity

Similar manipulations lead to expressions for transforming measurements of velocity from one system of coordinates to the other.

Henceforth, flydot derivative notation such as $\dot{x}$ will indicate a time derivative for a particular coordinate direction, e.g., $\dot{x} = u_x = \mathrm{d}x/\mathrm{d}t$ for the velocity $\mathbf{u} = (u_x, u_y, u_z)$.

First, taking the differential form of the time coordinate equation of the Lorentz transformation (which is for a primed time variable in terms unprimed variables) and dividing it through by $\mathrm{d}t$ yields

$$\frac{\mathrm{d}t'}{\mathrm{d}t} = \frac{1 - \dfrac{V}{c^2}\dfrac{\mathrm{d}x}{\mathrm{d}t}}{\sqrt{1 - V^2/c^2}} = \frac{1 - \dfrac{V\dot{x}}{c^2}}{\sqrt{1 - V^2/c^2}},$$

which connects the measurements $\mathrm{d}t'$ and $\mathrm{d}t$ of the time interval in the two systems S' and S between neighboring events which occur at neighboring points in space.

The spatial interval between the two events, when measured in system S, has as its x-component the distance which would be traveled with the component velocity $\dot{x}$ in the time $\mathrm{d}t$.

Next, with the previously obtained relation for $\dfrac{\mathrm{d}t'}{\mathrm{d}t}$ and the original formulation of the Lorentz transformation equations for primed variables in terms of unprimed variables, expressions for transforming measurements of velocity from one system of coordinates to the other may be obtained.

Differentiating the spatial coordinate equations for the Lorentz transformation with respect to t' and substituting the value for $\dfrac{\mathrm{d}t'}{\mathrm{d}t}$ yields

$$\begin{aligned} {u_x}' &= \frac{u_x - V}{1 - u_x V/c^2}, \\ {u_y}' &= \frac{u_y\sqrt{1 - V^2/c^2}}{1 - u_x V/c^2}, \\ {u_z}' &= \frac{u_z\sqrt{1 - V^2/c^2}}{1 - u_x V/c^2}. \end{aligned}$$

The physical interpretation of these transformation equations is that if for an observer in system S a point is found to be moving with the uniform velocity $\mathbf{u}$, its velocity $\mathbf{u}'$ as measured by an observer in system S' can be calculated from these equations.

It is often more convenient to have the transformation equations in the form in which they are solved for the unprimed quantities since this leads more readily to final expressions without the primes, thus the reciprocal equations to the velocity transformation equations previously obtained can be solved for the unprimed quantities in terms of the primed to yield

$$
\begin{aligned}
u_x &= \frac{u'_x + V}{1 + u'_x V/c^2}, \\
u_y &= \frac{u'_y\sqrt{1 - V^2/c^2}}{1 + u'_x V/c^2}, \\
u_z &= \frac{u'_z\sqrt{1 - V^2/c^2}}{1 + u'_x V/c^2}.
\end{aligned}
$$

6.2.6 Transformation Equation for the Lorentz Contraction Factor

Next, a transformation equation for the Lorentz contraction factor, i.e., the quantity $\sqrt{1 - V^2/c^2}$, for an object moving with the velocity $\mathbf{u}$ with respect to a given system of coordinates can be obtained from the velocity transformation equations previously obtained, namely,

$$
\sqrt{1 - u^2/c^2} = \frac{\sqrt{1 - u'^2/c^2}\sqrt{1 - V^2/c^2}}{1 + u_x{}'V/c^2}
$$

where $u^2 = u_x^2 + u_y^2 + u_z^2$.

6.2.7 The Lorentz Transformation Equations for Acceleration

Differentiating the velocity transformation equations yields the transformation equations for acceleration which can be written as

$$
\begin{aligned}
\dot{u}_x &= \left(1 + \frac{u_x{}'V}{c^2}\right)^{-3}\left(1 - \frac{V^2}{c^2}\right)^{3/2}\dot{u}_x{}', \\
\dot{u}_y &= \left(1 + \frac{u_x{}'V}{c^2}\right)^{-2}\left(1 - \frac{V^2}{c^2}\right)^{3/2}\dot{u}_y{}' - u'_y\frac{V}{c^2}\left(1 + \frac{u'_x V}{c^2}\right)^{-3}\left(1 - \frac{V^2}{c^2}\right)\dot{u}_x{}', \\
\dot{u}_z &= \left(1 + \frac{u_x{}'V}{c^2}\right)^{-2}\left(1 - \frac{V^2}{c^2}\right)^{3/2}\dot{u}_z{}' - u'_z\frac{V}{c^2}\left(1 + \frac{u'_x V}{c^2}\right)^{-3}\left(1 - \frac{V^2}{c^2}\right)\dot{u}_x{}'.
\end{aligned}
$$

6.2.8 Relativistic Conservation of Mass and Momentum for an Elastic Collision

Now that transformation equations have been obtained for position, the differential between neighboring events, velocity, and acceleration, proceed to consider conservation of mass and momentum in conjunction with these results.

Conservation of mass, i.e., that the total mass of a system of particles must remain constant over the course of their interactions with each other, may be written as

$$\sum m = \text{const.}$$

where the summation is taken over all of the particles in the system.

The components of the total momentum of the system in the x, y and z directions must also remain constant and this may be written

$$\sum mu_x = \text{const.},\ \sum mu_y = \text{const.},\ \sum mu_z = \text{const.},$$

where these sums are taken over the components of the momenta of all of the individual particles.

Note that these considerations only apply to particles which could interact by collision and do not account for systems where a continuous distribution of mass and momentum might have to be assigned to a field.

In Newtonian mechanics, one may use the simple Galilean transformation equations and these conservation statements can be satisfied in all systems of coordinates with the assumption that the mass of a particle is a constant independent of its velocity.

In relativistic mechanics, however, the Lorentz transformation equations must be used, and the statements of the conservation of mass and momentum shall be satisfied by allowing the mass of a particle to depend on its velocity.

To show that the mass of a particle must depend on its velocity, first consider the conservation of mass and momentum for the case of a very simple head-on collision between two similar elastic particles in two different systems of coordinates S' and S.

In the S' system, let the two particles be moving before collision with the velocities $+u'$ and $-u'$ parallel to the x-axis in such a way that a head-on collision can occur.

Since by hypothesis the two particles are perfectly similar and elastic, it is evident that they will first be brought to rest on collision and then rebound under the action of the resulting elastic forces, moving back over their original paths with the respective velocities $-u'$ and $+u'$ of the same magnitude as before but reversed in direction.

Now consider the same collision from the perspective of the second system of coordinates S while it is moving relative to the first system in the x-direction with the velocity $-V$.

In this second system of coordinates, denote the velocities of the two particles before the collision by u_1 and u_2, denote the masses of the two particles before the collision by m_1 and m_2, and denote by M the sum of the masses of the two particles at the exact instant in the course of the collision when they have come to relative rest.

At the instant of the collision when they have come to relative rest, both masses are moving with the velocity $+V$ with respect to the present system of coordinates, i.e., system S.

By virtue of the conservation of mass and momentum, which also must hold in this new set of coordinates, the total mass and total momentum of the two particles must be the same before the collision and at the instant of relative rest, thus

$$m_1 + m_2 = M,$$
$$m_1 u_1 + m_2 u_2 = MV.$$

Additionally, by the previously obtained transformation equations for velocity, the velocities u_1 and u_2 can be written in terms of their values $+u'$ and $-u'$ with respect to the original coordinates S, namely

$$u_1 = \frac{u' + V}{1 + u'V/c^2} \quad \text{and} \quad u_2 = \frac{-u' + V}{1 - u'V/c^2}.$$

Combining these equations and solving for the ratio of the two masses yields

$$\frac{m_1}{m_2} = \frac{1 + u'V/c^2}{1 - u'V/c^2},$$

which then due to the transformation equation for the Lorentz contraction factor yields

$$\frac{m_1}{m_2} = \frac{\sqrt{1 - u_2^2/c^2}}{\sqrt{1 - u_1^2/c^2}}.$$

Suppose that the two particles had the same initial mass, say m_0, when at rest.

The result is that the masses of the two particles become inversely proportional to $\sqrt{1 - u^2/c^2}$ when moving with the velocity u, thus

$$m = \frac{m_0}{\sqrt{1 - u^2/c^2}}$$

where m denotes the mass of a moving particle in terms of its velocity u and its rest mass m_0.

This suggests that the mass of a given particle will measure differently in different sets of coordinates since the velocity will be different in each respective set of coordinates.

Following from the transformation equation for the Lorentz contraction factor, the transformation of masses is obtained as

$$m = m' \frac{1 + u'_x V/c^2}{\sqrt{1 - V^2/c^2}}.$$

Differentiating with respect to time and simplifying yields a transformation equation for the rate at which the mass of a particle is changing due to change in velocity, namely

$$\frac{\mathrm{d}m}{\mathrm{d}t} = \frac{\mathrm{d}m'}{\mathrm{d}t'} + \frac{m'V}{c^2}(1 + u'_x V/c^2)^{-1} \frac{\mathrm{d}u'_x}{\mathrm{d}t'}.$$

6.2.9 Relativistic Force, Work, and Energy

It is no longer feasible as in Newtonian mechanics to define force as both mass times acceleration as well as the rate of change of momentum since the mass of a moving particle will change with its velocity.

Rather, force will be defined only as the rate of change of momentum since the principle of the equality of action and reaction forces is then equivalent to the principle of the conservation of momentum asserted earlier.

Hence the equation for the force **F** acting on a particle of mass m and velocity **u** is defined by the vector expression

$$\mathbf{F} = \frac{\mathrm{d}}{\mathrm{d}t}(m\mathbf{u}) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m_0\mathbf{u}}{\sqrt{1-u^2/c^2}}\right).$$

Note that from this definition in general force and acceleration will not be in the same direction as in Newtonian mechanics.

The previously derived transformation equations for position, velocity, and mass between coordinate systems can now be applied to obtain transformation equations for the components of force which can be written as

$$\begin{aligned}
F_x &= F'_x + \frac{u'_y V}{c^2 + u'_x V}F'_y + \frac{u'_z V}{c^2 + u'_x V}F_z, \\
F_y &= \frac{c^2\sqrt{1-V^2/c^2}}{c^2 + u'_x V}F'_y, \\
F_z &= \frac{c^2\sqrt{1-V^2/c^2}}{c^2 + u'_x V}F'_z.
\end{aligned}$$

Next, the work done on a particle by a force which displaces the particle through a distance in the direction of action and the resulting change in the kinetic energy of the particle will be considered.

Work due to such a force is given by

$$\mathrm{d}W = \mathbf{F}\cdot\mathrm{d}\mathbf{r}$$

where **r** is the radius vector determining the position of the particle.

The energy given to a particle by the action of a force is defined as equal to the work done on it.

By combining the definition for work with the vector expression for the force acting on a particle as the relativistic rate of change of momentum, the increase in kinetic energy can be written as

$$\begin{aligned}\mathrm{d}E &= m\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t}\cdot\mathrm{d}\mathbf{r} + \frac{\mathrm{d}m}{\mathrm{d}t}\mathbf{u}\cdot\mathrm{d}\mathbf{r}\\ &= m\mathbf{u}\cdot\mathrm{d}\mathbf{u} + \mathbf{u}\cdot\mathbf{u}\mathrm{d}m\\ &= mu\ \mathrm{d}u + u^2\ \mathrm{d}m.\end{aligned}$$

Substituting the previously obtained expression for mass as a function of velocity yields

$$\begin{aligned}\mathrm{d}E &= \frac{m_0 u}{(1-u^2/c^2)^{1/2}}\mathrm{d}u + \frac{m_0 u^3/c^2}{(1-u^2/c^2)^{3/2}}\mathrm{d}u\\ &= \frac{m_0 u\ \mathrm{d}u}{(1-u^2/c^2)^{3/2}}\end{aligned}$$

Integrating this expression from zero to u yields for the total kinetic energy of a particle of rest-mass m_0 moving with velocity u the expression

$$E = \frac{m_0 c^2}{\sqrt{1-u^2/c^2}} - m_0 c^2$$

which reduces at velocities small compared with that of light to the Newtonian case $E = \frac{1}{2}m_0 u^2$.

6.2.10 Einstein's Mass-Energy Equivalence

At this point, Einstein's fourth Annus Mirabilis paper [18] can be introduced to provide the foundation for the final step in the relativistic derivation of the momentum of a photon which is expected to agree with the expression for the same quantity given in Bohm's book.

This fourth 1905 paper by Einstein is titled *"Does the Inertia of a Body Depend Upon Its Energy-Content?"* and purportedly introduces the cliché $E = mc^2$, but as shall be seen shortly this is not exactly the case.

Einstein's mass-energy equivalence paper starts by considering a system of plane waves of light possessing the energy l in a system of coordinates given by (x, y, z) propagating at an angle ϕ with respect to the x-axis of the system.

Next, a new system of coordinates (ξ, η, ζ) is introduced moving in uniform parallel translation with respect to the original system (x, y, z) with its origin in motion along the x-axis with the velocity v.

The light as measured in the (ξ, η, ζ) system then possesses the energy

$$l^\star = l \frac{1 - \dfrac{v}{c}\cos\phi}{\sqrt{1 - v^2/c^2}}.$$

Einstein then considers a stationary body in the system (x, y, z) with an energy in that system referred to as E_0.

The energy of the same body but relative to the (ξ, η, ζ) system, moving again as previously with the velocity v, is referred to as H_0.

This body then sends out, in a direction at an angle ϕ with the x-axis, plane waves of light with energy $\frac{1}{2}L$ measured relative to (x, y, z), and simultaneously an equal quantity of light in the the opposite direction.

Meanwhile the body remains at rest with respect to the (x, y, z) system.

Energy must be conserved in this process with respect to both systems of coordinates.

The energy of the body after the emission of light is then denoted E_1 or H_1 measured relatively to the (x, y, z) or (ξ, η, ζ) system, respectively.

Then, by using his first relation for $l^\star$, Einstein obtained that

$$\begin{aligned} E_0 &= E_1 + \frac{1}{2}L + \frac{1}{2}L, \\ H_0 &= H_1 + \frac{L}{\sqrt{1 - v^2/c^2}}. \end{aligned}$$

Combining these then yielded

$$H_0 - E_0 - (H_1 - E_1) = L\left\{\frac{1}{\sqrt{1 - v^2/c^2}} - 1\right\}.$$

Since H and E are energy values of the same body referred to from different coordinate systems which are in motion relative to each other, and the body is at rest in one of the two systems (i.e., system (x, y, z)), then the difference $H - E$ can differ from the kinetic energy K of the body with respect to the other system (i.e., system (ξ, η, ζ)) only by an additive constant C.

This constant depends on the choice of the arbitrary additive constants of the energies H and E, thus

$$\begin{aligned} H_0 - E_0 &= K_0 + C, \\ H_1 - E_1 &= K_1 + C \end{aligned}$$

since C does not change during the emission of light.

This implies that

$$K_0 - K_1 = L\left\{\frac{1}{\sqrt{1 - v^2/c^2}} - 1\right\}.$$

Einstein omitted the details in his paper, but note that in order to make the next step which is effectively the punchline of the paper, the Lorentz factor in the right hand side of this result for the difference $K_0 - K_1$ is Taylor series expanded and then higher order terms are neglected.

To see this, denote the Lorentz factor as $\gamma = \dfrac{1}{\sqrt{1 - v^2/c^2}}$.

To expand this function, utilize the Generalized Binomial Theorem, namely

$$(1 + x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha - 1)}{2!}x^2 + \frac{\alpha(\alpha - 1)(\alpha - 2)}{3!}x^3 + \cdots$$

where $\alpha = -1/2$ and $x = \dfrac{v^2}{c^2}$.

Hence the first few terms of the expansion is

$$\Rightarrow \gamma = 1 + \frac{1}{2}\left(\frac{v}{c}\right)^2 + \frac{3}{8}\left(\frac{v}{c}\right)^4 + \frac{5}{16}\left(\frac{v}{c}\right)^6 + \mathcal{O}\left(\frac{v}{c}\right)^8$$

Neglecting magnitudes of fourth and higher orders and substituting the truncated expansion back into the expression for $K_0 - K_1$ yields

$$K_0 - K_1 = L\{\gamma - 1\} = L\{1 + \frac{1}{2}\left(\frac{v}{c}\right)^2 - 1\} = \frac{1}{2}\frac{L}{c^2}v^2.$$

Einstein didn't actually write $E = mc^2$, but instead said that *if a body gives off the energy* L *in the form of radiation, its mass diminishes by* L/c^2 [18].

In other words, a quantity of energy E, which Einstein originally denoted L, always has immediately associated with it a mass m of the amount $m = \dfrac{E}{c^2}$.

6.2.11 Relativistic Momentum of a Photon[5]

This implies that considering the momentum associated with the transfer of energy with the velocity $\mathbf{u}$, one obtains

$$\mathbf{p} = m\mathbf{u} = \frac{E}{c^2}\mathbf{u}.$$

In the case of a photon traveling at the speed of light c in the direction of a unit propagation vector $\hat{\mathbf{k}}$, then

$$\mathbf{u} = c\hat{\mathbf{k}} \Rightarrow \mathbf{p} = \frac{E}{c}\hat{\mathbf{k}} \tag{2}$$

which is the same as the expression (1) from Bohm's book.

---

[5] Follows from section 27 of [46].

7 Quantization of Photon Energy

Continuing with the *"Quantum Theory"* content [1, Ch. 2, pp. 32–33], consider how this radiation momentum is affected by the quantization of energy.

Since the energy comes in discrete units of hv, the momentum comes in units of hv/c which can be manipulated to yield

$$\mathbf{p} = \frac{hv}{c}\hat{\mathbf{k}} = \frac{h}{\lambda}\hat{\mathbf{k}} = \hbar\mathbf{k}$$

where $\hbar = h/2\pi$.

Therefore, when a radiation oscillator is excited to its n^{th} quantum state, it has energy $E = nhv$ and momentum $\mathbf{p} = n\hbar\mathbf{k}$ for integer n.

The interpretation of this is that its energy and momentum behave like a collection of n particles, each with energy hv and momentum $\hbar\mathbf{k}$.

The state of excitation of a radiation field can therefore be specified via the number of equivalent particles corresponding to each $\mathbf{k}$.

As has been discussed, the electromagnetic field can interact with matter only by the emission or absorption of full quanta.

Similarly, the interactions between matter and light take place only by means of the emission or absorption of full photons.

References

[1] David Bohm. *Quantum Theory.* Prentice-Hall, Englewood Cliffs, NJ, 1951.

[2] Christiaan Huygens. *Traité de la lumière.* Pieter van der Aa, Leiden, 1690.

[3] Isaac Newton. *Opticks: Or, A Treatise of the Reflexions, Refractions, Inflexions and Colours of Light.* Samuel Smith and Benjamin Walford, London, 1704. Also Two Treatises of the Species and Magnitude of Curvilinear Figures.

[4] Augustin-Jean Fresnel. *De la Lumière.* Académie des sciences, Paris, 1822.

[5] Augustin-Jean Fresnel. *Oeuvres complètes d'Augustin Fresnel*, volume 2. Imprimerie Impériale, Paris, 1868.

[6] James Clerk Maxwell. A dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London*, 155:459–512, 1865.

[7] Heinrich Hertz. Ueber die ausbreitungsgeschwindigkeit der electrodynamischen wirkungen. *Annalen der Physik*, 270(7):551–569, 1888.

[8] Max Planck. Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum. *Verhandlungen der Deutschen Physikalischen Gesellschaft*, 2:237–245, 1900. Presented December 14, 1900.

[9] Wilhelm Wien. On the division of energy in the emission-spectrum of a black body. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 43(262):214–220, 1897. Translated from *Annalen der Physik*, vol. 58, pp. 662–669 (1896).

[10] Albert Einstein. Die plancksche theorie der strahlung und die theorie der spezifischen wärme. *Annalen der Physik*, 327(1):180–190, 1907.

[11] P. Debye. Zur theorie der spezifischen wärmen. *Annalen der Physik*, 344(14):789–839, 1912.

[12] Max Planck and Morton Masius. *The Theory of Heat Radiation.* P. Blakiston's Son & Co., Philadelphia, PA, 1914. Page 135.

[13] Marco Giliberti and Luisa Lovisetti. *Old Quantum Theory and Early Quantum Mechanics: A Historical Perspective Commented for the Inquiring Reader.* Challenges in Physics Education. Springer, 2024. Page 159.

[14] Albert Einstein. Über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt. *Annalen der Physik*, 322(6):132–148, 1905.

[15] Albert Einstein. On a heuristic point of view about the creation and conversion of light, 1905. English translation on Wikisource, accessed 18 May 2026.

[16] Albert Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik*, 322(8):549–560, 1905.

[17] Albert Einstein. Zur elektrodynamik bewegter körper. *Annalen der Physik*, 322(10):891–921, 1905.

[18] Albert Einstein. Ist die trägheit eines körpers von seinem energieinhalt abhängig? *Annalen der Physik*, 323(13):639–641, 1905.

[19] Heinrich Hertz. Ueber einen Einfluss des ultravioletten Lichtes auf die electrische Entladung. *Annalen der Physik*, 267(8):983–1000, 1887.

[20] François Rabelais. *Gargantua and Pantagruel*. Penguin Classics, London, 2006.

[21] Gaspar Schott. *Mechanica hydraulico-pneumatica*. Sumptu Haeredum Joannis Godefridi Schönwetteri, excudebat Henricus Pigrin, Würzburg, 1657. Accessit experimentum novum Magdeburgicum.

[22] Otto von Guericke. *Experimenta Nova (ut vocantur) Magdeburgica de Vacuo Spatio*. Joannem Jansonium a Waesberge, Amsterdam, 1672. Quibus accesserunt simul certa quædam alia experimenta.

[23] Michael Faraday. VIII. Experimental researches in electricity. — Thirteenth series. *Philosophical Transactions of the Royal Society of London*, 128:125–168, 1838.

[24] Julius Plücker. Ueber die Einwirkung des Magneten auf die elektrischen Entladungen in verdünnten Gasen. *Annalen der Physik und Chemie*, 179(1):88–106, 1858.

[25] William Crookes. The Bakerian Lecture.—On the illumination of lines of molecular pressure, and the trajectory of molecules. *Philosophical Transactions of the Royal Society of London*, 170:135–164, 1879.

[26] William Crookes. On radiant matter. *Nature*, 20(512):419–423, 1879.

[27] Joseph John Thomson. Cathode Rays. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 44(269):293–316, 1897.

[28] Philipp Lenard. Ueber die electrostatischen Eigenschaften der Kathodenstrahlen. *Annalen der Physik und Chemie*, 300(2):279–289, 1898.

[29] Philipp Lenard. Ueber die lichtelektrische Wirkung. *Annalen der Physik*, 313(5):149–198, 1902.

[30] Max Planck. Ueber das Gesetz der Energieverteilung im Normalspektrum. *Annalen der Physik*, 309(3):553–563, 1901.

[31] George Gabriel Stokes. On the change of refrangibility of light. *Philosophical Transactions of the Royal Society of London*, 142:463–562, 1852.

[32] Robert Andrews Millikan. The Isolation of an Ion, a Precision Measurement of its Charge, and the Correction of Stokes's Law. *Science*, 32(822):436–448, 1010.

[33] Robert Andrews Millikan. On the Elementary Electrical Charge and the Avogadro Constant. *Physical Review*, 2(2):109–143, 1913.

[34] Robert Andrews Millikan. A Direct Determination of "h". *Physical Review*, 4(1):73–75, 1914.

[35] Robert Andrews Millikan. A Direct Photoelectric Determination of Planck's "h". *Physical Review*, 7(3):355–388, 1916.

[36] Nobel Prize Outreach. The Nobel Prize in Physics 1921. `https://www.nobelprize.org/prizes/physics/1921/summary/`, 2026. Accessed: 2026-05-19.

[37] Nobel Prize Outreach. The Nobel Prize in Physics 1923. `https://www.nobelprize.org/prizes/physics/1923/summary/`, 2026. Accessed: 2026-05-19.

[38] Thomas Young. The Bakerian Lecture: Experiments and calculations relative to physical optics. *Philosophical Transactions of the Royal Society of London*, 94:1–16, 1804.

[39] Paul Adrien Maurice Dirac. *The Principles of Quantum Mechanics.* International Series of Monographs on Physics. Clarendon Press, Oxford, 1930.

[40] Ernest O. Lawrence and Jesse W. Beams. The Element of Time in the Photoelectric Effect. *Physical Review*, 32(3):478–485, 1928.

[41] Robert d'Escourt Atkinson. On the emission of light hydrogen atoms. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 116(773):81–103, 1927.

[42] Marcus Ossiander, Johannes Riemensberger, Stefan Neppl, Michael Mittermair, Martin Schäffer, Armin Duensing, Max S. Wagner, Reiner Heider, Maximilian Wurzer, Christian Gerl, Matthias Bothschafter, Johannes V. Barth, Reinhard Kienberger, Ralph Ernstorfer, Vladislav S. Yakovlev, René Pazourek, Stefan Nagele, Joachim Burgdörfer, Ferenc Krausz, and Florian Feist. Absolute timing of the photoelectric effect. *Nature*, 561(7723):374–377, 2018.

[43] David Bohm. A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I. *Physical Review*, 85(2):166–179, 1952.

[44] David Bohm. A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. II. *Physical Review*, 85(2):180–193, 1952.

[45] Albert Einstein. Die Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften*, 1915(2):844–847, 1915. Session of 25 November 1915.

[46] Richard Chace Tolman. *Relativity, Thermodynamics, and Cosmology.* International Series of Monographs on Physics. Clarendon Press, Oxford, 1934.

[47] Hendrik Antoon Lorentz. Electromagnetic Phenomena in a System Moving with Any Velocity Smaller than That of Light. *Proceedings of the Royal Netherlands Academy of Arts and Sciences*, 6:809–831, 1904.