

Einstein's Original Paper

Published in Annalen der Physik **17** (1905) 132–148

(Translator D. TER HAAR)

(Highlights in bold and comments in color by Vinay Agarwala)

On a Heuristic Point of View about the Creation and Conversion of Light

There exists an essential formal difference between the theoretical pictures physicists have drawn of gases and other ponderable bodies and Maxwell's theory of electromagnetic processes in so-called empty space. Whereas we assume the state of a body to be completely determined by the positions and velocities of an, albeit very large, still finite number of atoms and electrons, we use for the determination of the electromagnetic state in space continuous spatial functions, so that a finite number of variables cannot be considered to be sufficient to fix completely the electromagnetic state in space. **According to Maxwell's theory, the energy must be considered to be a continuous function in space for all purely electromagnetic phenomena, thus also for light, while according to the present-day ideas of physicists the energy of a ponderable body can be written as a sum over the atoms and electrons.** The energy of a ponderable body cannot be split into arbitrarily many, arbitrarily small parts, while the energy of a light ray, emitted by a point source of light is according to Maxwell's theory (or in general according to any wave theory) of light distributed continuously over an ever increasing volume.

The wave theory of light which operates with continuous functions in space has been excellently justified for the representation of purely optical phenomena and it is unlikely ever to be replaced by another theory. One should, however, bear in mind that optical observations refer to time averages and not to instantaneous values and notwithstanding the complete experimental verification of the theory of diffraction, reflection, refraction, dispersion, and so on, **it is quite conceivable that a theory of light involving the use of continuous functions in space will lead to contradictions with experience, if it is applied to the phenomena of the creation and conversion of light.**

In fact, it seems to me that the observations on "black-body radiation", photoluminescence, the production of cathode rays by ultraviolet light and other phenomena involving the emission or conversion of light can be better understood on the assumption that the energy of light is distributed discontinuously in space. **According to the assumption considered here, when a light ray starting from a point is propagated, the energy is not continuously distributed over an ever increasing volume, but it consists of a finite number of energy quanta, localised in space, which move without being divided and which can be absorbed or emitted only as a whole.**

In the following, I shall communicate the train of thought and the facts which led me to this conclusion, in the hope that the point of view to be given may turn out to be useful for some research workers in their investigations.

Maxwell's theory expresses the energy of the electromagnetic radiation as a continuous function in space. On the other hand, the kinetic energy of particles of gas in space is discontinuous and treated through statistical functions.

1. On a Difficulty in the Theory of "Black-body Radiation"

To begin with, we take the point of view of Maxwell's theory and electron theory and consider the following case. Let there be in a volume completely surrounded by reflecting walls, a number of gas molecules and electrons moving freely and exerting upon one another conservative forces when they approach each other that is, colliding with one another as gas molecules according to the kinetic theory of gases. [This assumption is equivalent to the proposition that the average kinetic energies of gas molecules and electrons are equal to one another in temperature equilibrium. It is well known that Mr. Drude has theoretically derived in this way the relation between the thermal and electrical conductivities of metals.] Let there further be a number of electrons which are bound to points in space, which are far from one another, by forces proportional to the distance from those points and in the direction towards those points. These electrons are also assumed to be interacting conservatively with the free molecules and electrons as soon as the latter come close to them. ***We call the electrons bound to points in space "resonators"; they emit and absorb electromagnetic waves with definite periods.***

According to present-day ideas on the emission of light, ***the radiation in the volume considered***—which can be found for the case of dynamic equilibrium on the basis of the Maxwell theory—***must be identical with the "black-body radiation"***—at least provided we assume that resonators are present of all frequencies to be considered.

For the time being, we neglect the radiation emitted and absorbed by the resonators and ***look for the condition for dynamic equilibrium corresponding to the interaction (collisions) between molecules and electrons.*** Kinetic gas theory gives for this the condition that the average kinetic energy of a resonator electron must equal the average kinetic energy corresponding to the translational motion of a gas molecule. If we decompose the motion of a resonator electron into three mutually perpendicular directions of oscillation, ***we find for the average value E_{av} of the energy of such a linear oscillatory motion***

$$E_{av} = (R/N) T,$$

where **R** is the gas constant, **N** the number of "real molecules" in a gramme equivalent and **T** the absolute temperature. This follows as the energy E_{av} is equal to $\frac{2}{3}$ of the kinetic energy of a free molecules of a monatomic gas since the time averages of the kinetic and the potential energy of a

resonator are equal to one another. If, for some reason—in our case because of radiation effects—one manages to make the time average of a resonator larger or smaller than E_{av} , collisions with the free electrons and molecules will lead to an energy transfer to or from the gas which has a non-vanishing average. **Thus, for the case considered by us, dynamic equilibrium will be possible only, if each resonator has the average energy E_{av} .**

We can now use a similar argument for the interaction between the resonators and the radiation which is present in space. Mr. Planck has derived for this case the condition for dynamic equilibrium under the assumption that one can consider the radiation as the most random process imaginable. He found

$$E_v = (L^3/8\pi v^2)\rho_v$$

where E_v is the average energy of a resonator with eigen-frequency v (per oscillating component), L the velocity of light, v the frequency and $\rho_v dv$ the energy per unit volume of that part of the radiation which has frequencies between v and $v + dv$. If the radiation energy of frequency v is not to be either decreased or increased steadily, we must have

$$(R/N) T = E_{av} = E_v = (L^3/8\pi v^2)\rho_v$$

$$\rho_v = (R/N) (8\pi v^2 / L^3) T$$

This relation, which we found as the condition for dynamic equilibrium does not only lack agreement with experiment, but it also shows that in our picture there can be no question of a definite distribution of energy between aether and matter. The greater we choose the range of frequencies of the resonators, the greater becomes the radiation energy in space and in the limit we get

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

The classical treatment of energy as a continuous function in space fails to predict the energy spectrum observed for the black body radiation.

2. On Planck's Determination of Elementary Quanta

We shall show in the following that determination of elementary quanta given by Mr. Planck is, to a certain extent, independent of the theory of "black-body radiation" constructed by him.

Planck's formula for ρ_v which agrees with all experiments up to the present is

$$\rho_v = \frac{\alpha v^3}{e^{\beta v/T} - 1},$$

where $\alpha = 6 \cdot 10 \times 10^{-56}$, $\beta = 4.866 \times 10^{-11}$.

For large values of T/v , that is, for long wavelengths and high radiation densities, this formula has the following limiting form

$$\rho_v = \frac{\alpha}{\beta} v^2 T.$$

One sees that this formula agrees with the one derived in section 1 from Maxwell theory and electron theory. By equating the Coefficients in the two formulae, we get

$$\frac{R}{N L^3} = \frac{\alpha}{\beta}$$

or

$$N = \frac{\beta}{\alpha} \frac{8\pi R}{L^3} = 6.17 \times 10^{23},$$

that is, one hydrogen atom weighs $1/N = 1.62 \times 10^{-24}$ g. This is exactly the value found by Mr. Planck, which agrees satisfactorily with values of this quantity found by different means.

We thus reach the conclusion : ***the higher the energy density and the longer the wavelengths of radiation, the more usable is the theoretical basis used by us; for short wavelengths and low radiation densities, however, the basis fails completely.***

In the following, we shall consider "black-body radiation", basing ourselves upon experience without using a picture of the creation and propagation of the radiation.

Max Planck's postulate, "energy is proportional to the frequency of radiation" predicts the energy spectrum observed for the black body radiation. Planck's equation provides the classical results for low frequencies, showing that radiation energy is a continuous function in space at low frequencies only.

3. On the Entropy of the Radiation

The following considerations are contained in a famous paper by Mr. W. Wien and are only mentioned here for the sake of completeness.

Consider radiation which takes up a volume V . We assume that the observable properties of this radiation are completely determined if we give the radiation energy $\rho(v)$ for all frequencies. [This is an arbitrary assumption. Of course, one sticks to this simplest, assumption until experiments force us to give it up.] As we may assume that radiations of different frequencies can be separated without work or heat, we can write the entropy of the radiation in the form

$$S = v \int_0^{\infty} \phi(\rho, v) dv,$$

where φ is a function of the variables ρ and ν . One can reduce φ to a function of one variable only by formulating the statement that the entropy of radiation between reflecting walls is not changed by an adiabatic compression. We do not want to go into this, but at once investigate how one can obtain the function φ from the radiation law of a black body.

In the case of "black-body radiation", ρ is such a function of ν or that the entropy is a maximum for a given energy, that is,

$$\delta \int_0^{\infty} \phi(\rho, \nu) d\nu = 0, \quad \text{if} \quad \delta \int_0^{\infty} \rho d\nu = 0.$$

From this it follows that for any choice of $\delta\rho$ as function of ν

$$\int_0^{\infty} \left(\frac{\partial\phi}{\partial\rho} - \lambda \right) \delta\rho d\nu = 0,$$

where λ is independent of ν . In the case of black-body radiation, $\partial\phi/\partial\rho$ is thus independent of ν .

If the temperature of a black-body radiation in a volume $v = 1$ increases by dT , we have the equation

$$dS = \int_{\nu=0}^{\nu=\infty} \frac{\partial\phi}{\partial\rho} d\rho d\nu,$$

or, as $\partial\phi/\partial\rho$ is independent of ν :

$$dS = \frac{\partial\phi}{\partial\rho} dE.$$

As dE is equal to the heat transferred and as the process is reversible, we have also

$$dS = \frac{1}{T} dE.$$

Through comparing, we get

$$\frac{\partial\phi}{\partial\rho} = \frac{1}{T^*}$$

This is the black-body radiation law. One can thus from the function φ obtain the black-body radiation law and conversely from the latter the function φ , through integration, bearing in mind that φ vanishes for $\rho = 0$.

Einstein assumes that the observable properties of black-body radiation are completely determined when the radiation energy is given for all frequencies. This allows him to form an entropy function of energy density that is independent of frequency.

4. Limiting Law for the Entropy of Monochromatic Radiation for Low Radiation Density

From the observations made so far on “black-body radiation” it is clear that the law

$$\rho = \alpha v^3 e^{-\beta v/T}$$

put forward originally for “black-body radiation” by Mr. W. Wien is not exactly valid. However; for large values of v/T , it is in complete agreement with experiment. We shall base our calculations on this formula, though bearing in mind that our results are valid only within certain limits.

First of all, we get from this equation

$$\frac{1}{T} = -\frac{1}{\beta v} \ln \frac{\rho}{\alpha v^3},$$

and then, if we use the relation found in the preceding section

$$\phi(\rho, v) = -\frac{\rho}{\beta v} \left[\ln \frac{\rho}{\alpha v^3} - 1 \right].$$

Let there now be radiation of energy E with a frequency between v and $v + dv$ and let the volume of the radiation be V . The entropy of this radiation is

$$S = v\phi(\rho, v) dv = -\frac{E}{\beta v} \left[\ln \frac{E}{v\alpha v^3 dv} - 1 \right].$$

If we restrict ourselves to investigating the dependence of the entropy on the volume occupied by the radiation, and if we denote the entropy of the radiation by S_0 if it occupies a volume V_0 , we get

$$S - S_0 = \frac{E}{\beta v} \ln \frac{v}{v_0}.$$

This equation shows that the entropy of a monochromatic radiation of sufficiently small density varies with volume according to the same rules as the entropy of a perfect gas or of a dilute solution. The equation just found will in the following be interpreted on the basis of the principle, introduced by Mr. Boltzmann into physics, according to which the entropy of a system is a function of the probability of its state.

Einstein relates his entropy function to Wein’s proposal, which accurately describes the black-body spectrum at high frequencies. Einstein is then able to show that the entropy of monochromatic radiation follows the same rules as the entropy of a perfect gas.

Thus Einstein proves that the energy distribution of radiation becomes particle-like at high frequencies. This is an ingenious way of arriving at this conclusion.

5. Molecular-Theoretical Investigation of the Volume-dependence of the Entropy of Gases and Dilute Solutions

When calculating the entropy in molecular gas theory one often uses the word “probability” in a sense which is not the same as the definition of probability given in probability theory. Especially, often “cases of equal probability” are fixed by hypothesis under circumstances where the theoretical model used is sufficiently definite to deduce probabilities rather than fixing them by hypothesis. ***I shall show in a separate paper that when considering thermal phenomena it is completely sufficient to use the so-called “statistical probability”, and I hope thus to do away with a logical difficulty which is hampering the consistent application of Boltzmann’s principle.*** At the moment, however, I shall give its general formulation and the application to very special cases.

If it makes sense to talk about the probability of a state of a system and if, furthermore, any increase of entropy can be considered as a transition to a more probable state, the entropy S_1 of a system will be a function of the probability W_1 of its instantaneous state. If, therefore, one has two systems which do not interact with one another, one can write

$$S_1 = \phi_1(W_1), \quad S_2 = \phi_2(W_2).$$

If one considers these two systems as a single system of entropy S and probability W we have

$$S = S_1 + S_2 = \phi(W) \quad \text{and} \quad W = W_1 \cdot W_2.$$

This last relation states that the states of the two systems are independent. From these equations it follows that

$$\phi(W_1 \cdot W_2) = \phi_1(W_1) + \phi_2(W_2),$$

and hence finally $\phi_1(W_1) = C \ln W_1 + \text{const},$

$$\phi_2(W_2) = C \ln W_2 + \text{const},$$

$$\phi(W) = C \ln W + \text{const}.$$

The quantity C is thus a universal constant; it follows from kinetic gas theory that it has the value R/N where the constants R and N have the same meaning as above. If S_0 is the entropy of a certain initial state of the system considered and W the relative probability of a state with entropy S , we have in general

$$S - S_0 = \frac{R}{N} \ln W.$$

We now consider the following special case. Let us consider a number, n , moving points (e.g., molecules) in a volume V_0 . Apart from those, there may be in this space arbitrarily many other moving points of some kind or other. ***We do not make any assumptions about the laws according to which the points considered move in space, except that as far as their motion is concerned no part of space—and no direction—is preferred above others. The number of the (first-mentioned) points which we are considering be moreover so small that we can neglect their mutual interaction.***

There corresponds a certain entropy S_0 to the system under consideration, which may be, for instance, a perfect gas or a dilute solution. Consider now the case where a part V of the volume V_0 contains all n moving points while otherwise nothing is changed in the system. This state clearly corresponds to a different value, S_1 of the entropy, and we shall now use Boltzmann's principle to determine the entropy difference.

We ask : how large is the probability of this state relative to the original state? Or: how large is the probability that at an arbitrary moment all n points moving independently of one another in a given volume V_0 are (accidentally) in the volume V ?

One gets clearly for this probability, which is a "statistical probability";

$$W = \left(\frac{v}{v_0}\right)^n;$$

one obtains from this, applying Boltzmann's principle:

$$S - S_0 = R \frac{n}{N} \ln \frac{v}{v_0}.$$

It must be noted that it is unnecessary to make any assumptions about the laws, according to which the molecules move, to derive this equation from which one can easily derive thermodynamically the Boyle-Gay-Lussac law and the same law for the osmotic pressure.

Einstein is demonstrating that the use of "statistical probability", when applied to a large number of discrete particles, is compatible with physical laws.

6. Interpretation of the Expression for the Volume-dependence of the Entropy of Monochromatic Radiation according to Boltzmann's Principle

In Section 4, we found for the volume-dependence of the entropy of monochromatic radiation the expression

$$S - S_0 = \frac{E}{\beta v} \ln \frac{v}{v_0}.$$

If we write this equation in the form

$$S - S_0 = \frac{R}{N} \ln \left[\left(\frac{v}{v_0} \right)^{NE/R\beta v} \right],$$

and compare it with the general formula which expresses Boltzmann's principle,

$$S - S_0 = \frac{R}{N} \ln W,$$

we arrive at the following conclusion :

If monochromatic radiation of frequency ν and energy E is enclosed (by reflecting walls) in a volume V_0 , the probability that at an arbitrary time the total radiation energy is in a part V of the volume V_0 will be

$$W = \left(\frac{V}{V_0}\right)^{NE/R\beta\nu}.$$

From this we then conclude:

Monochromatic radiation of low density behaves—as long as Wien’s radiation formula is valid—in a thermodynamic sense, as if it consisted of mutually independent energy quanta of magnitude $R\beta\nu/N$.

We now wish to compare the average magnitude of the “blackbody” energy quanta with the average kinetic energy of the translational motion of a molecule at the same temperature. The latter is $(3/2)RT/N$, while we get from Wien’s formula for the average magnitude of the energy quantum :

$$\frac{\int_0^{\infty} \alpha \nu^3 e^{-\beta\nu/T} d\nu}{\int_0^{\infty} \frac{N}{R\beta\nu} \alpha \nu^3 e^{-\beta\nu/T} d\nu} = 3 \frac{R}{N} T.$$

If monochromatic radiation—of sufficiently low density—behaves, as far as the volume-dependence of its entropy is concerned, as a discontinuous medium consisting of energy quanta of magnitude $R\beta\nu/N$, it is plausible to investigate whether the laws on creation and transformation of light are also such as if light consisted of such energy quanta. This question will be considered in the following.

Einstein is determining the magnitude of energy quanta when the density is sufficiently low for monochromatic radiation.

7. On Stokes’ Rule

Consider monochromatic light which is changed by photoluminescence to light of a different frequency; in accordance with the result we have just obtained, we assume that both the original and the changed light consist of energy quanta of magnitude $(R/N)\beta\nu$, where ν is the corresponding frequency. We must then interpret the transformation process as follows. Each initial energy quantum of frequency ν_1 is absorbed and is—at least when the distribution density of the initial energy quanta is sufficiently low—by itself responsible for the creation of a light quantum of frequency ν_2 ; possibly in the absorption of the initial light quantum at the same time also light quanta of frequencies ν_3, ν_4, \dots as well as energy of a different kind (e.g. heat) may be generated. It is immaterial through what intermediate processes the final result is brought about. Unless we can consider the photo-luminescing substance as a continuous source of energy, the energy of a final

light quantum can, according to the energy conservation law, not be larger than that of an initial light quantum; we must thus have the condition

$$\frac{R}{N}\beta v_2 \leq \frac{R}{N}\beta v_1, \text{ or } v_2 \leq v_1$$

This is the well-known Stokes' rule.

We must emphasize that according to our ideas the intensity of light produced must—other things being equal—be proportional to the incident light intensity for weak illumination, as every initial quantum will cause one elementary process of the kind indicated above, independent of the action of the other incident energy quanta. Especially, there will be no lower limit for the intensity of the incident light below which the light would be unable to produce photoluminescence.

According to the above ideas about the phenomena deviations from Stokes' rule are imaginable in the following cases:

1. When the number of the energy quanta per unit volume involved in transformations is so large that an energy quantum of the light produced may obtain its energy from several initial energy quanta.
2. When the initial (or final) light energetically does not have the properties characteristic for "black-body radiation" according to Wien's law; for instance, when the initial light is produced by a body of so high a temperature that Wien's law no longer holds for the wavelengths considered.

This last possibility needs particular attention. According to the ideas developed here, it is not excluded that a "non-Wienian radiation", even highly-diluted, behaves energetically differently than a "black-body radiation" in the region where Wien's law is valid.

Einstein is using the new idea of "energy quanta" to explain the Stoke's Rule for photoluminescence.

8. On the Production of Cathode Rays by Illumination of Solids

The usual idea that the energy of light is continuously distributed over the space through which it travels meets with especially great difficulties when one tries to explain photo-electric phenomena, as was shown in the pioneering paper by Mr. Lenard.

According to the idea that the incident light consists of energy quanta with an energy $R\beta v/N$, one can picture the production of cathode rays by light as follows. Energy quanta penetrate into a surface layer of the body, and their energy is at least partly transformed into electron kinetic energy. The simplest picture is that a light quantum transfers all of its energy to a single electron; we shall assume that that happens. We must, however, not exclude the possibility that electrons only receive part of the energy from light quanta. An electron obtaining kinetic energy inside the body will have

lost part of its kinetic energy when it has reached the surface. Moreover, we must assume that each electron on leaving the body must produce work P , which is characteristic for the body. Electrons which are excited at the surface and at right angles to it will leave the body with the greatest normal velocity.

The kinetic energy of such electrons is

$$\frac{R}{N}\beta v - P$$

If the body is charged to a positive potential Π and surrounded by zero potential conductors, and if Π is just able to prevent the loss of electricity by the body, we must have

$$\Pi \epsilon = \frac{R}{N}\beta v - P,$$

where ϵ is the electrical mass of the electron, or

$$\Pi E = R\beta v - P',$$

where E is the charge of a gram equivalent of a single-valued ion and P' is the potential of that amount of negative electricity with respect to the body. [If one assumes that it takes a certain amount of work to free a single electron by light from a neutral molecule, one has no need to change this relation; one only must consider P' to be the sum of two terms.]

If we put $E = 9.6 \times 10^3$, $\Pi \times 10^8$ is the potential in Volts which the body assumes when it is irradiated in a vacuum.

To see now whether the relation derived here agrees, as to order of magnitude, with experiments, we put $P' = 0$, $v = 1.03 \times 10^{15}$ (corresponding to the ultraviolet limit of the solar spectrum) and $\beta = 4.866 \times 10^{-11}$. We obtain $\Pi \times 10^7 = 4.3$ Volt, a result which agrees, as to order of magnitude, with Mr. Lenard's results.

If the formula derived here is correct, Π must be, if drawn in Cartesian coordinates as a function of the frequency of the incident light, a straight line, the slope of which is independent of the nature of the substance studied.

As far as I can see, our ideas are not in contradiction to the properties of the photoelectric action observed by Mr. Lenard. ***If every energy quantum of the incident light transfers its energy to electrons independently of all other quanta, the velocity distribution of the electrons, that is, the quality of the resulting cathode radiation, will be independent of the intensity of the incident light; on the other hand, ceteris paribus [other things being equal], the number of electrons leaving the body should be proportional to the intensity of the incident light.***

As far as the necessary limitations of these rules are concerned, we could make remarks similar to those about the necessary deviations from the Stokes rule.

In the preceding, we assumed that the energy of at least part of the energy quanta of the incident light was always transferred completely to a single electron. If one does not make this obvious assumption, one obtains instead of the earlier equation the following one

$$\Pi E + P' \leq R\beta v.$$

For cathode-luminescence, which is the inverse process of the one just considered, we get by a similar argument

$$\Pi E + P' \geq R\beta v.$$

For the substances investigated by Mr. Lenard, ΠE is always considerably larger than $R\beta v$, as the voltage which the cathode rays must traverse to produce even visible light is, in some cases a few hundred, in other cases thousands of volts. We must thus assume that the kinetic energy of an electron is used to produce many light energy quanta.

Einstein is brilliantly verifying the calculated value of "energy quanta" from the experimental data obtained from the study of photoelectricity.

9. On the Ionization of Gases by Ultraviolet Light

We must assume that when a gas is ionized by ultraviolet light, always one absorbed light energy quantum is used to ionize just one gas molecule. From this follows first of all that the ionization energy (that is, the energy theoretically necessary for the ionization) of a molecule cannot be larger than the energy of an effective, absorbed light energy quantum. If J denotes the (theoretical) ionization energy per gram equivalent, we must have

$$R\beta v \geq J.$$

According to Lenard's measurements, the largest effective wavelength for air is about 1.9×10^{-5} cm, or

$$R\beta v = 6.4 \times 10^{12} \text{ erg} \geq J.$$

An upper limit for the ionization energy can also be obtained from ionization voltages in dilute gases. According to J. Stark the smallest measured ionization voltage (for platinum anodes) in air is about **10 Volt**. [In the interior of the gas, the ionization voltage for negative ions is anyhow five times larger.] We have thus an upper limit of 9.6×10^{12} for J which is about equal to the observed one. There is still another consequence, the verification of which by experiment seems to me to be very important. If each light energy quantum which is absorbed ionizes a molecule, the following relation should exist between the absorbed light intensity L and the number j of moles ionized by this light:

$$j = \frac{L}{R\beta v}.$$

This relation should, if our ideas correspond to reality, be valid for any gas which—for the corresponding frequency—does not show an appreciable absorption which is not accompanied by ionization.

Einstein is testing his ideas to explain the existing experimental observations, and proving the viability of the idea of “energy quantum” or “light quantum”.