plosives, which increasingly threaten all mankind.

Any appraisal of Einstein must acknowledge his greatness as a man as well as his greatness as a physicist. He was the supreme humanist, standing on the side of justice for all men regardless of the cost to himself. He believed deeply in the worth and dignity of every human being, and he fought against the evils that men everywhere are called to suffer at the hands of those who misuse power. One of his biographers has remarked that Einstein was a man of the sort who appears but once in a century—as a physicist it would be more realistic to say once in two centuries, for this was the approximate time that separated Newton and Einstein.

No man, with the possible exception of Newton, has so influenced the science of an era as Einstein influenced the physics of the twentieth century. His papers, whose full significance for the structure of the atom only became evident many years after they were published, are models of excellence and simplicity, with much of the development based a good deal more on physical reasoning than on mathematical formulation. Although at first sight the relationship between his work and the course of atomism is not always evident, a closer analysis shows that much of what we accept today about atoms would be untenable if any of Einstein's basic discoveries were discarded.

Thus, without his concept of the photon as an unchanging entity, such phenomena as the photoelectric effect, the photoionization of atoms, fluorescence, photoluminescence, could not be understood. Without the special theory of relativity such things as nuclear energy, the electron spin, the fine structure of spectral lines, positrons, and antimatter, in general, could not be understood. Today physicists are beginning to feel, more and more strongly, that a complete picture of the structure of matter is possible only if Einstein's general theory of relativity, published fifty years ago, is properly incorporated into atomic theory.

Before we go into a detailed discussion of Einstein's paper on the photon, we shall sketch briefly the chronology of his work.

He began with a series of papers, between 1902 and 1904, in which he developed the essential features of statistical mechanics, without foreknowledge of the similar work that Willard Gibbs had already completed in 1901. In 1905, he applied his statistical mechanics to the analysis of Brownian motion, using the kinetic theory of gases. In the same year and in the same journal, the *Annalen der Physik,* two more of Einstein's fundamental papers appeared. In one, he developed the theory of the photon, which we shall discuss presently, and used it to explain the photoelectric effect, Stokes's law, and photoluminescence, among other phenomena. In the second, he began his series of famous papers on the theory of relativity. In the ensuing ten years, Einstein presented his theory of specific heats, in which, for the first time, the quantum theory was made to account for the observed behavior of the specific heats of solids as the absolute temperature approaches zero. In this same creative period, he published his proof that energy and mass are equivalent, deriving the famous equation $E = mc^2$. He also published an analysis indicating that the equivalence of gravitational and inertial mass is not a mere accident of nature, but the basis of a profound physical principle that leads to a new theory of gravity. He also extended the application of statistical mechanics to all physical systems and showed that the relationship of statistical mechanics to thermodynamics is valid under the most general conditions. Shortly thereafter, he published his general theory of relativity, his general statistical derivation of Planck's law of radiation, in which only atomic processes of absorption and emission are assumed, and his first paper on cosmology, which ushered in modern cosmology. Finally, he recognized the importance of the wave properties of particles and the need to consider these properties in the statistical mechanics of such particles.

The remarkable thing about all Einstein's investigations is that they are all of a fundamental nature. In some cases, they opened up totally new realms in science. His papers demonstrated an uncanny ability to penetrate to the heart of the most obscure problem. There seems never to have been any doubt in his mind that he had the correct answers, even when confronting the most profound questions that disturbed his contemporaries. His principal concern, it appears, was to supply the answers in an understandable manner, with as little formalism as possible. All of these papers are marked by bold departures from the accepted paths and by confident applications of new and untried ideas. Thus, in his papers on the specific heats of solids, he departs immediately from the idea, accepted until then, that the atomic vibrators composing these solids obey the ordinary classical laws of mechanics; instead he assumes that they are governed by the quantum theory. This was, indeed, a revolutionary step of the first magnitude, since it had previously been thought that quantum effects were to be ascribed to radiation only. Because Einstein obtained more nearly correct results for the specific heats of solids with the quantum theory (impossible with classical physics), physicists realized that the quantum theory would have to be applied to all atomic processes. The next great advance in this direction was made by Bohr in his monumental work on atomic spectra.



Had Einstein restricted himself to any one of the fields listed above, his contributions would still have marked him as one of the great physicists of our time. Although most people suppose that he was awarded the Nobel Prize for his theory of relativity, the prize was actually given for his paper of photons and the explanation of the photoelectric effect. The same paper, "Concerning a Heuristic Point of View about the Creation and Transformation of Light," is reproduced at the end of this commentary. It is remarkable because, although it deals with one of the most puzzling problems that physicists had to cope with, it is amazingly free of complex mathematical formulas. Indeed, except for one or two equations in the entire paper, the analysis is carried out with elementary algebra that can be followed by a good third-year high-school student. This paper is also of special interest because it shows that statistical mechanics, which up to then had been applied only to systems of particles, can also be applied to radiation in a container. Such radiation acts in many respects like a perfect gas. By carrying out the analogy between the statistics of molecules in a gas and the statistics of radiation in a container, Einstein established the existence of photons as unchanging entities under all conditions.

Although Planck had introduced the quantum of action to account for the spectral distribution of black-body radiation, the concept of the photon as a real entity was not a very popular one and was rejected by most physicists at the time. Planck himself felt very uncomfortable about the photon, and considered it more or less as a useful device to derive the correct radiation equations. He was inclined to picture the photon as having reality (if it had any at all) only during the processes of absorption and emission. At all other times, then, radiation had only a wave structure and character. Einstein departed completely from this tentative position that sought a compromise between classical physics and the new quantum hypothesis. He went over entirely to the quantum theory. He states his position very early in the paper that we are discussing and accepts the concept in the following words: "It appears to me, in fact, that the observations . . . can be understood better on the assumption that the energy in light is distributed discontinuously in space." Having stated this revolutionary position unequivocally, Einstein demonstrated the correctness of his assumption by applying to a container of black-body radiation, at the absolute temperature $T$, the statistical laws of Boltzmann and Maxwell, and those that he himself had developed in connection with the behavior of systems of particles. He carries through the analysis in two broad steps. He first considers freely moving particles, electrons and molecules, bound resonators, and harmonic oscillators in the light of the classical laws of thermodynamics and statistical mechanics, intermingled with the radiation in the container and interacting with it, as well as interacting among themselves. All are in dynamic equilibrium. He treats this ensemble of radiation, particles, and resonators according to the classical electromagnetic theory of Maxwell—the continuous distribution of energy in black-body radiation—and the classical theory of the electron. Therefore, since thermal equilibrium exists in this ensemble (the temperature is constant), it follows (from the classical laws of thermodynamics and statistical mechanics) that the mean kinetic energy of the resonators must equal that of the free particles, the molecules and electrons. The kinetic energy of the free particles is given by the equation $E = \frac{3}{2}kT$, ($k$ being the Boltzmann constant) and hence is proportional to the absolute temperature $T$. In this derivation there is implicit the concept that each oscillator can have all possible energies consistent with the condition that it be in equilibrium with the freely moving molecules at the given temperature. By equilibrium we mean that, on the average, each molecule and each resonator has the same energy.

Einstein next considers the interaction of the *radiation* and the resonators when equilibrium exists; he assumes that Maxwell's theory governs this interaction. Therefore, the resonators can absorb and emit radiant energy continuously, that is, classically. Under these conditions the mean energy of a resonator of frequency $\nu$ must be directly proportional to the energy density of the radiation of that frequency and inversely proportional to the square of the frequency itself.

On equating this result for the mean energy of a resonator with that found by the methods of the previous paragraph, he finds that the energy density of the radiation in the container is proportional to the absolute temperature and the square of the frequency of the radiation.

Einstein points out that this is clearly incorrect, since the density of radiant energy in the higher frequencies would grow larger and larger if the frequency were continually increased, which is contrary to the observed facts. It is clear, then, that the idea of a continuous distribution of radiant energy in the container, in accordance with Maxwell's theory, is incorrect, although, as Einstein shows later in the paper, this is approximately true for very high temperatures and long wavelengths, that is, for large radiation densities.

To show that this result for the long wavelengths becomes more and more accurate with increasing absolute temperature, Einstein starts out from the correct Planck formula for the radiation density and considers it for large values of the absolute temperature divided by the frequency of



the radiation. Under these conditions he shows that Planck's formula is the same as the one developed from the classical theory of Maxwell, that is, it is proportional to the square of the frequency and the absolute temperature. This leads to two conclusions: first, that the classical theory of radiation (that is, the wave theory with the continuous distribution of energy) is valid if we are dealing with dense quantities of radiation of long wavelength; second, the constants appearing in Planck's formula must be related to the gas constant and Avogadro's number since these constants appear in the classical formula. In fact, by equating the classical formula for the density of the radiation to that obtained from the Planck formula for high temperatures and long wavelengths, Einstein derives the mass of the hydrogen atom, that is, the reciprocal of Avogadro's number. This, to Einstein, was a clear indication that Planck's formula is correct and that the classical formula is only a correct approximation under certain conditions.

In the next part of the paper, Einstein considers the more important question of the nature of radiation when the density of the radiation is small and when the wavelength is very short. Under these conditions, one may replace Planck's formula by a somewhat simpler one, which was first introduced by Wien but which, in Einstein's words, "is completely satisfied experimentally for large values of $\nu/T$" (with $\nu$ the frequency and $T$ the temperature).

Einstein now compares the behavior of the radiation in this case with that of a perfect gas by introducing the entropy of the radiation which he can easily calculate from the Wien formula. He then shows that the entropy of the radiation expressed in terms of the volume of the radiation is of exactly the same mathematical form as the formula for the entropy of a perfect gas. To obtain the entropy of the gas in the form that is suitable for comparison with the entropy of the radiation, Einstein applies to the gas an important principle that was first stated by Boltzmann, viz.: that the entropy of a system is related to the probability for the state of the system. In so doing, Einstein gave a definite physical meaning to Boltzmann's equation relating entropy and probability; he defined the probability of a state as the length of time during which the system remains in this state, relative to some standard time.

In form, the equation Einstein derived for the entropy of an ideal gas is identical, in its dependence on volume, with the formula he derived for the entropy of monochromatic radiation in a container. Therefore, he concludes that the chance of finding, at any given moment, in some smaller volume of the container its total monochromatic radiation is expressed in a formula that is the same in form as the probability that the molecules in a perfect gas, distributed throughout a volume $v_0$, all will be found in a smaller volume $v$ at any given moment.

Einstein's formula for the entropy of a dilute, short-wavelength, monochromatic radiation gas of frequency $\nu$ leads, via Boltzmann's relationship between probability and entropy, to the probability $(v/v_0)^{E/h\nu}$. This is the probability for finding all the radiation in the smaller volume $v_0$, where $E$ is the energy of the radiation and $h$ is Planck's constant of action. Since the probability for finding all the molecules, $n$, of a perfect gas in the volume $v_0$ at any time is $(v/v_0)^n$, Einstein concludes that the two exponents, $E/h\nu$ and $n$, in these two formulas must be the same. Since $n$ is the number of molecules in the gas, it follows from this that $E/h\nu$ is the number of distinct particles (quanta or photons) of radiation in the radiation. Since $E$ is the total energy of the radiation, it follows that each quantum or particle of radiation has an amount of energy $h\nu$.

This is exactly the content of Planck's quantum hypothesis, but Einstein's results go further than Planck's hypothesis in that they show that radiation always consists of quanta or photons. Einstein states in his paper that follows:

> *Monochromatic radiation of small energy density (within the validity range of the Wien radiation formula) behaves in thermodynamic theoretical relationships as though it consisted of distinct independent* energy quanta of magnitude $h\nu$—[Editors' italics].

In the last part of the paper Einstein applies these conclusions to the explanation of two effects: Stokes's rule and the photoelectric effect. We shall consider here only the second of these two phenomena. It had been known for a number of years before Einstein's work, as the result of Heinrich Hertz's experiments and observations, that when light strikes a metal surface electrons are emitted. This is known as the photoelectric effect. The energy of the emitted electrons does not depend on the intensity of the light used, but only on the color of the light. This cannot be understood on the basis of Maxwell's classical electromagnetic theory. According to this theory, the intensity of the beam incident on the metal surface should determine the energy of the ejected electrons, which is proportional to the square of the speed with which each electron comes off the surface.

Einstein explained the phenomenon very easily, however, by means of his photon hypothesis (as will be seen in his paper) in the following words:

> According to the concept that the exciting radiation consists of energy quanta with energy content $h\nu$, the production of cathode rays by light can be understood as follows. Quanta of energy penetrate into the surface layer of the body and their energy, at least in part, is transformed into kinetic energy of electrons. The simplest explanation is that a quantum transfers all its energy to a single electron. . . . If each quantum of energy of the exciting light gives up its energy to an electron independently of all the other quanta, then the velocity distribution of the electrons, that is, the quality of the produced cathode ray, is independent of the intensity of the exciting radiation; on the other hand, the number of electrons leaving the body, all other conditions being the same, will depend on the intensity of the exciting radiation.



During the next three years Einstein came back to the problem of the nature of radiation and, in a series of papers as brilliant and simple as the one discussed above, established the existence of the photon beyond any doubt.

At the same time, he showed that the statistical methods that worked so well in the analysis of the Brownian motion can be applied with the same success to radiation. He used these statistical methods to analyze the statistical fluctuations of radiation in a volume. If one considers a small part of a large volume containing radiation, the energy of the radiation in this small volume fluctuates from moment to moment. Einstein calculated this fluctuation using Planck's formula for the energy density of the radiation.

He showed that if $E$ is the mean energy of the radiation of frequency $\nu$ in a volume, then the square of the fluctuation of this energy is equal to

$$h\nu\left(\frac{E}{V}\right)+\left(\frac{8\pi\nu^2d\nu}{c^3}\right)^{-1}\left(\frac{E}{V}\right)^2$$

where $d\nu$ is the frequency range and $V$ is the total volume of the radiation. This is a very remarkable result, as Einstein first pointed out, for it can be shown that although the second term $\left(\frac{8\pi\nu^2d\nu}{c^3}\right)^{-1}\left(\frac{E}{V}\right)^2$ in this expression can be derived from classical electromagnetic theory, the first term, $h\nu\,(E/V)$, cannot be so derived and can only be accounted for by assuming that radiation has particle properties.

Einstein had thus indicated mathematically when energy can be expected to behave as a group of particles and when as a wave. The two terms in this expression show that the wave properties are dominant when the frequency is small (long wavelengths) and the energy density $E/V$ is large. On the other hand, the first term (hence, the particle properties of radiation) dominates when $E/V$ is small (dilute radiation) and the frequency is large (short wavelengths).

Not content with this analysis alone, Einstein went on to consider a small mirror suspended freely in black-body radiation and analyzed its motion as though it were a Brownian particle in a gas. As the mirror is bombarded by photons from all sides, it fluctuates back and forth and carries out a Brownian motion. Again Einstein found that this motion consists of two parts, one of which can be explained by means of Maxwell's classical electrodynamic theory. But the other effect can only be understood if one accepts the quantum hypothesis.

## *EINSTEIN*

# Concerning a Heuristic Point of View about the Creation and Transformation of Light [1]

THERE IS A PROFOUND FORMAL difference between the theoretical representations of gases and other ponderable bodies which physicists have constructed and Maxwell's theory of electromagnetic processes in so-called empty space. Whereas we may consider the state of a body as being completely determined by the positions and velocities of, to be sure, a very large but finite number of atoms and electrons, we must use continuous, spatial functions to specify the electromagnetic state of a region, so that a finite number of parameters cannot be considered as sufficient to describe completely the electromagnetic state of a region of space. According to Maxwell's theory in all cases of pure electromagnetic phenomena, hence in the case of light, the energy must be considered as a continuous spatial function, whereas the energy of a ponderable body, according to the current concepts of physicists, can be represented by a sum taken over the atoms and electrons. The energy of a ponderable body can break up into arbitrarily many, arbitrarily small parts, whereas the energy of a ray of light emitted by a point source of light distributes itself continuously throughout an ever-increasing volume of space according to Maxwell's theory (or, more generally, according to any wave theory).

The wave theory, operating with continuous spatial functions, has proved to be correct in representing purely optical phenomena and will probably not be replaced by any other theory. One must, however, keep in mind that the optical observations are concerned with temporal mean values and not with instantaneous values, and it is possible, in spite of the complete experimental verification of the theory of diffraction, reflection, refraction, dispersion, and so on, that the theory of light that operates with continuous spatial functions may lead to contradictions with observations if we apply it to the phenomena of the generation and transformation of light.

[1] Albert Einstein, trans. Editors, *Annalen der Physik,* 17 (1905), 132–148.



It appears to me, in fact, that the observations on "black-body radiation," photoluminescence, the generating of cathode rays with ultraviolet radiation, and other groups of phenomena related to the generation and transformation of light can be understood better on the assumption that the energy in light is distributed discontinuously in space. According to the presently proposed assumption the energy in a beam of light emanating from a point source is not distributed continuously over larger and larger volumes of space but consists of a finite number of energy quanta, localized at points of space, which move without subdividing and which are absorbed and emitted only as units.

In what follows, I want to present the thinking and indicate the facts that have led me along the present path in the hope that the point of view associated with these ideas may prove useful to some researchers in their investigations.

## I. CONCERNING CERTAIN DIFFICULTIES IN THE THEORY OF BLACK-BODY RADIATION

We begin by adopting the point of view of Maxwell's theory and the electron theory and consider the following case. We suppose that a volume of space completely enclosed by perfectly reflecting walls contains a number of gas molecules and electrons, which move about freely and exert conservative forces on each other when they get close to each other, that is, they can collide with one another like gas molecules according to the kinetic theory of gases. Further let a number of electrons be anchored at widely separated points by forces which are directed to these points and which are proportional to the displacements from the points. These electrons also are to be in conservative interaction with the free electrons and molecules when these free particles get very close to them. We call the anchored electrons "resonators"; they emit and absorb electromagnetic waves of definite period.

According to the current point of view about the emission of light, the radiation in the enclosed volume, which for the case of dynamical equilibrium is present on the basis of Maxwell's theory, must be identical with the "black-body radiation," at least when resonators of all frequencies that are to be considered are present.

We disregard for the moment the radiation that is absorbed and emitted by the resonators and investigate the condition for dynamical equilibrium for interactions [collisions] among the gas molecules and the electrons. The kinetic theory of gases gives as this condition that the mean energy of a resonator electron must equal the mean kinetic energy of the translational motion of a gas molecule. If we resolve the motion of a resonator electron into three mutually perpendicular components, we find for the mean value of the energy of such a linear vibratory motion

$$\overline{E} = R/N\,T$$

where $R$ is the gas constant, $N$ is the number of real molecules in a gram equivalent, and $T$ is the absolute temperature. Because of the equality of the time average of the kinetic and potential energy of a resonator, the energy $\overline{E}$ is ⅔ as large as the kinetic energy of a monatomic gas molecule. If through some cause—in our case through radiative processes—the mean energy of the resonator were larger than $E$, the collisions with the free electrons and gas molecules would lead, on the average, to a transfer to, or an absorption from, the gas of a non-zero amount of energy. We see then that in our case dynamical equilibrium is possible only when each resonator has a mean energy $\overline{E}$.

We carry through a similar analysis relative to the interaction of the resonators and the radiation in the container. Planck has derived the condition for dynamical equilibrium in this case under the assumption that the radiation can be considered as the most random phenomenon possible. He found

$$\overline{E}_\nu = \frac{L^3}{8\pi\nu^2}\,\rho_\nu;$$

$\overline{E}_\nu$ is here the mean energy of a resonator of proper frequency $\nu$ (per oscillatory component or degree of freedom), $L$ the speed of light, $\nu$ the frequency, and $\rho_\nu d\nu$ the energy per unit volume of the radiation in the frequency range from $\nu$ to $\nu + d\nu$.

If the radiation energy of frequency $\nu$ is neither to increase nor to decrease continuously on the whole, we must have

$$\frac{R}{N}T = \overline{E} = \overline{E}_\nu = \frac{L^3}{8\pi\nu^2}\cdot\rho_\nu$$

$$\rho_\nu = \frac{R}{N}\cdot\frac{8\pi\nu^2}{L^3}\cdot T.$$

This relationship obtained as the condition for dynamical equilibrium not only does not agree with observation, but it also states that in our picture there can be no talk of a definite distribution of energy between ether and matter. The larger the frequency range that is chosen for the resonators, the larger is the radiation energy in the container, and in the limit we obtain

$$\int_0^\infty \rho_\nu d\nu = \frac{R}{N}\cdot\frac{8\pi}{L^3}\cdot T\int_0^\infty \nu^2 d\nu = \infty$$



## II. PLANCK'S DERIVATION OF THE ELEMENTARY QUANTA

We wish to demonstrate in what follows that Planck's derivation of the elementary quanta is to a certain degree independent of his theory of "black radiation."

Planck's formula for $\rho_\nu$ which satisfies all the conditions required up to now reads

$$\rho_\nu = \frac{\alpha\nu^3}{e^{\frac{\beta\nu}{T}} - 1}$$

where

$$\alpha = 6.10 \times 10^{-56} = \frac{8\pi h}{L^3}$$

$$\beta = 4.866 \times 10^{-11} = \left(\frac{h}{k}\right).$$

For large values of $T/\nu$, that is, long wavelengths and large radiation densities, this formula in the limit goes over into the following one

$$\rho_\nu = \frac{\alpha}{\beta}\nu^2 T$$

We see that this formula is the same as the one in Section I developed from the Maxwell and the electron theory. By equating the coefficients in both formulae we obtain

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

or

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

that is, a hydrogen atom has a mass of $\frac{1}{N}$ grams $= 1.62 \times 10^{-24}$ grams. This is precisely the value found by Planck, which agrees very satisfactorily with the value of this constant found in other ways.

We thus arrive at the following conclusion: The larger the energy density and the wavelength of any radiation, the more applicable are the theoretical principles we have used; for small wavelengths and small energy densities, however, these principles break down completely.

In the following we shall consider the "black radiation" and its observed characteristics without introducing a model for the emission and propagation of radiation.

## III. CONCERNING THE ENTROPY OF RADIATION

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

Let us consider radiation which occupies a volume $V$. We assume that the observable characteristics of this radiation are fully known when the energy density $\rho(\nu)$ is given for all frequencies. Since radiations of different frequencies may be considered as separable from each other without doing work or introducing heat, the entropy of the radiation may be written in the form

$$S = V\int_0^{\infty} \phi(\rho,\nu)\,d\nu$$

where $\phi$ is a function of the variables $\rho$ and $\nu$. $\phi$ can be reduced to a function of only one variable by expressing in formulae the statement that the entropy of radiation is unaltered by compressing it adiabatically between perfectly reflecting walls. However, we do not want to go into this here but instead consider at once how the function $\phi$ is to be determined from the radiation law of black bodies.

For "black radiation" $\rho$ is such a function of $\nu$ that the entropy is a maximum for a given energy, that is, that

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

if

$$\delta\int_0^{\infty} \rho\,d\nu = 0,$$

From this it follows that for every choice of $\delta\rho$ as a function of $\nu$

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

where $\lambda$ is independent of $\nu$. Thus for "black radiation" $\frac{\partial\phi}{\partial\rho}$ is independent of $\nu$.



For a temperature increase $dT$ of black radiation of volume $V = 1$ we have the equation

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

or, since $\frac{\partial\phi}{\partial\rho}$ is independent of $\nu$,

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

Since $dE$ is equal to the heat supplied to the radiation and the process is reversible, we have

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

By comparing these two formulae we find

$$\frac{\partial\phi}{\partial\rho} = \frac{1}{T}.$$

This is the law of black radiation. We thus can obtain the law of black radiation from the function $\phi$ and conversely if we know this law we can find $\phi$ by integration, keeping in mind that $\phi$ vanishes if $\rho = 0$.

## IV. LIMITING LAW OF THE ENTROPY OF MONOCHROMATIC RADIATION FOR SMALL RADIATION DENSITY

From the considerations up to now about "black radiation," it follows that the original law proposed by W. Wien

$$\rho = \alpha\nu^3 e^{-\frac{\beta\nu}{T}}$$

is not precisely correct. However, it is satisfied experimentally for large values of $\frac{\nu}{T}$. We shall base our analysis on this formula but keep in mind that our results are valid only within certain limits.

From this formula we first see that

$$\frac{1}{T} = -\frac{1}{\beta\nu}\log\frac{\rho}{\alpha\nu^3}$$

and further, by taking into account the relationship obtained in the previous paragraph

$$\phi(\rho,\nu) = -\frac{\rho}{\beta\nu}\left\{\log\frac{\rho}{\alpha\nu^3} - 1\right\}.$$

We consider now radiation of energy $E$ whose frequency lies between $\nu$ and $\nu + d\nu$. Let this radiation occupy the volume $V$. The entropy of this radiation is

$$S = V\phi(\rho,\nu)\,d\nu = -\frac{E}{\beta\nu}\left\{\log\frac{E}{V\alpha\nu^3 d\nu} - 1\right\}.$$

If we limit ourselves to considering the dependence of the entropy on the volume of the radiation and designate the entropy of the radiation as $S_0$ when its volume is $V_0$, we obtain

$$S - S_0 = \frac{E}{\beta\nu}\log\left(\frac{V}{V_0}\right).$$

This equation shows that the entropy of monochromatic radiation of sufficiently small density varies with volume like the entropy of an ideal gas or like a very dilute solution. The relationship just obtained will be interpreted in what follows on the basis of the principle introduced into physics by Boltzmann, according to which the entropy of a system is a function of the probability of its state.

## V. MOLECULAR THEORETIC INVESTIGATION INTO THE DEPENDENCE OF THE ENTROPY OF GASES AND DILUTE SOLUTIONS ON THEIR VOLUMES

In calculating the entropy by molecular theoretic methods the word "probability" is often given a meaning which does not correspond to the definition of probability as it is given in the calculus of probability.

In particular, "cases of equal probability" are often determined hypothetically in situations where the applied theoretical pictures are definite enough to permit a deduction instead of the hypothetical method. I shall demonstrate in a separate paper that in dealing with thermal processes we can manage completely with the so-called "statistical probability" and hope thereby to overcome a logical difficulty which still stands in the way of carrying out the Boltzmann principle. Here, however, we shall concern ourselves only with its general formulation and its application to quite special cases.



If there is any meaning attached to talking of the probability of the state of a system, if, further, each entropy increase is to be taken as a transition to a more probable state, then the entropy $S$, of a system is a function of the probability $W$, of the momentary state. If, then, there are two noninteracting systems $S_1$ and $S_2$, we can set

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

If we consider these two systems as forming a single system of entropy $S$ and probability $W$, we have

$$S = S_1 + S_2 = \phi(W)$$

and

$$W = W_1 W_2.$$

The last relationship states that the states of the two systems are phenomena that are independent of each other.

From these equations we obtain

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

and from this, finally

$$\phi_1(W_1) = C\log W_1 + \text{const.}$$
$$\phi_2(W_2) = C\log W_2 + \text{const.}$$
$$\phi(W) = C\log W + \text{const.}$$

The constant $C$ is thus a universal constant; from the kinetic theory of gases its value is found to be $R/N$ where $R$ and $N$ have the same meanings as above. If $S_0$ is the entropy of a given system in its initial state and $W$ is the relative probability in a state of entropy $S$, we have in general

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

We treat first the following special case. In a volume $V_0$ let there be a number ($n$) of moving points (for example, molecules upon which we shall carry out our considerations). Besides these, there may also be present other moving points of any arbitrary kind. We shall assume nothing about the law that governs the motion of these points except to specify that the motion shall be such that no region of space (nor direction) shall be preferred above any other. The number of specified (the first mentioned) moving points shall further be so small that we may neglect the interaction of one of these points with another.

This system, which, for example, may be an ideal gas or a dilute solution has an entropy $S_0$. We consider a part $V$ of the volume $V_0$ and suppose that all $n$ moving points are transferred to the volume $V$ without any other change in the system. This state of the system obviously has a different value of the entropy ($S$) and we shall now calculate the entropy difference by means of Boltzmann's principle.

We ask: How large is the probability of this last considered state relative to the initial state? Or, how large is the probability that at some particular moment all $n$ independently moving particles in the given volume $V_0$ are found (by chance) in the volume $V$.

For this probability, which is a statistical probability, we find clearly the value

$$W = \left(\frac{V}{V_0}\right)^n .$$

We obtain from this by applying the Boltzmann principle

$$S - S_0 = R\left(\frac{n}{N}\right)\log\left(\frac{V}{V_0}\right) .$$

It is remarkable that for the derivation of this equation, from which the law of Boyle–Gay-Lussac and the equivalent law of osmotic pressure can be easily derived by thermodynamics, we had to make no assumption about the law governing the motion of the molecules.

## VI. INTERPRETATION OF THE EXPRESSION FOR THE DEPENDENCE OF THE ENTROPY OF MONOCHROMATIC RADIATION ON THE VOLUME ACCORDING TO THE BOLTZMANN PRINCIPLE

In Section IV we obtained the expression

$$S - S_0 = \frac{E}{\beta\nu}\log\left(\frac{V}{V_0}\right)$$

for the dependence of the entropy of monochromatic radiation on its volume. If we write this formula in the form

$$S - S_0 = \frac{R}{N}\log\left[\left(\frac{V}{V_0}\right)^{\frac{N}{R}\frac{E}{\beta\nu}}\right]$$



and compare it with the general formula expressing Boltzmann's principle

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

we reach the following conclusion:

If monochromatic radiation of frequency $\nu$ and energy $E$ is enclosed in the volume $V_0$ (by perfectly reflecting walls), then the probability that at any moment all the radiation energy will be found in the partial volume $V$ of the volume $V_0$ is given by

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

From this we conclude further:

Monochromatic radiation of small energy density (within the validity range of the Wien radiation formula) behaves in thermodynamic theoretical relationships as though it consisted of distinct independent energy quanta of magnitude $R\beta\frac{\nu}{N}$.

We wish further to compare the mean energy of the energy quanta of "black radiation" with the mean kinetic energy of the center of mass motion of a molecule for a given temperature. The latter is $\frac{3}{2}(R/N)T$, whereas using Wien's formula, we obtain for the mean energy of the energy-quanta

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

If then as far as the dependence of entropy on volume goes, monochromatic radiation (of sufficiently small density) behaves like a discontinuous medium consisting of energy-quanta of magnitude $R\beta\nu/N$, it is reasonable to inquire if the laws of emission and transformation of light are so constituted as though the light were composed of these same energy-quanta. We shall concern ourselves with this question in the next section.

## VII. ON STOKES'S RULE

Let monochromatic light be transformed by photoluminescence into light of a different frequency and, in accordance with the result just derived, let the exciting as well as the excited radiation consist of energy quanta of magnitude $\left(\frac{R}{N}\right)\beta\nu$, where $\nu$ is the frequency of the radiation. The transformation process is then to be understood as follows:

Each exciting quantum of frequency $\nu_1$ is absorbed and gives—at least for sufficiently small density distribution of the exciting energy quanta—by itself an opportunity for the emission of a quantum of frequency $\nu_2$; eventually it is possible after the absorption of the exciting quantum also for quanta of frequencies $\nu_3$, $\nu_4$, and so on, as well as other kinds of energy (e.g., heat) to be emitted simultaneously. It is not essential for these end results to determine what the intermediate processes are. If we are not to look upon the photoluminescing material as an unending source of energy, then according to the energy principle, the energy of an excited quantum cannot exceed the energy of the exciting light quantum; the following relationship must therefore be valid

$$\frac{R}{N}\beta\nu_2 \leqq \frac{R}{N}\beta\nu_1$$

or

$$\nu_2 \leqq \nu_1$$

This is the well-known Stokes's law.

In particular one must emphasize that for weak irradiation, the emitted amount of light, all other things being equal, must be proportional to the intensity of the exciting light since each exciting energy quantum generates an elementary process of the kind described above, independently of the action of the other exciting energy quanta. In particular there is no lower limit for the intensity of the exciting light below which the light is unable to excite light emission.

Departures from Stokes's law are conceivable according to the above picture of the process in the following cases:

1. if the number per unit volume of simultaneously transformed energy-quanta is so large that an energy-quantum of the excited light can receive its energy from many exciting energy quanta.

2. if the exciting (or excited) light is not of the same energy constitution as that of "black radiation" in the range where Wien's law is valid; if, for example, the exciting light is emitted by a body of so high a temperature that the pertinent wavelengths are no longer governed by Wien's law.

This last possibility deserves special interest. According to the concepts developed here one may not exclude the possibility that a "non-Wien type of radiation" may differ in its energetic behavior from that of "black radiation" in the region where Wien's law is valid.



## VIII. ON THE PRODUCTION OF CATHODE RAYS BY IRRADIATING SOLID BODIES

The traditional view that the energy of light is distributed continuously through the region illuminated by the light runs into great difficulty in trying to explain photoelectric phenomena, as was outlined in a trail-blazing paper by Lenard.

According to the concept that the exciting radiation consists of energy quanta with energy content $\left(\frac{R}{N}\right)\beta\nu$, the production of cathode rays by light can be understood as follows: Quanta of energy penetrate into the surface layer of the body and their energy, at least in part is transformed into kinetic energy of electrons. The simplest explanation is that a quantum transfers all its energy to a single electron; we shall assume that this occurs. We shall, however, not exclude the possibility that electrons can absorb only parts of the energy of light quanta. An interior electron with kinetic energy will have lost some of this kinetic energy by the time it reaches the surface. Besides this we must assume that each electron will have to do some work (an amount characteristic of the body) when it leaves the body. The electrons lying right at the surface of the body will leave the body with the greatest velocity normal to the surface. The kinetic energy of such electrons is

$$\frac{R}{N}\beta\nu - P.$$

If the body is charged to the positive potential $\Pi$ and surrounded by conductors at zero potential, and if $\Pi$ is large enough to prevent a discharge of the body, then we must have

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

Where $\epsilon$ is the electric charge of the electron, or

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

where $E$ is the charge of a gram equivalent of a single charged ion and $P'$ is the potential of this amount of negative charge relative to the body.

If we place $E = 9.6 \times 10^3$, then $\Pi \cdot 10^{-8}$ is the potential in volts that the body acquires on being irradiated in a vacuum.

In order to see at first if the derived relationship is of the right order of magnitude as obtained empirically we place $P' = 0$, $\nu = 1.03 \times 10^{15}$ (corresponding to the ultraviolet limit of the solar spectrum) and $\beta = 4.866 \times 10^{-11}$, we obtain $\Pi \cdot 10^7 = 4.3$ volts which agrees in order of magnitude with the results of Lenard.

If the derived formula is correct, then $\pi$ must be a linear function of the frequency whose slope depends on the nature of the material being studied.

Our point of view, as far as I can see, does not contradict Lenard's observed properties of the photoelectric phenomena. If each quantum of energy of the exciting light gives up its energy to an electron independently of all the other quanta, then the velocity distribution of the electrons, that is, the characteristic of the produced cathode ray, is independent of the intensity of the exciting radiation; on the other hand the number of electrons leaving the body, all other conditions being the same, will depend on the intensity of the exciting radiation.

Concerning the probable domain of validity of the above laws we may make the same observations as for Stokes's rule.

In the above we assumed that the energy of at least a part of the quantum of the exciting light is given completely to one electron. If we do *not* make this reasonable assumption, we obtain in place of the above equation the following:

$$\Pi E + P' \leqq R\beta\nu.$$

For cathode luminescence, which is the inverse of the process discussed above, we obtain by analogous reasoning

$$\Pi E + P' \geqq R\beta\nu.$$

For the substances investigated by Lenard $PE$ is significantly larger than $R\beta\nu$ since the potential through which the cathode rays had to move, just to emit visible light, equaled hundreds of volts in some cases, and thousands of volts in other cases. We may thus assume that the kinetic energy of an electron is used to produce numerous light quanta.

## IX. ON THE IONIZATION OF GASES WITH ULTRAVIOLET LIGHT

We assume that in the ionization of a gas by ultraviolet light, one quantum of energy ionizes just one molecule. From this it follows, first of all, that the ionization energy (that is, the theoretical work needed to ionize) of a molecule cannot exceed the energy of the absorbed quantum that is effective in ionizing the molecule. If $J$ is the (theoretical) ionization energy per gram equivalent, we must have

$$R\beta\nu \geqq J.$$



But according to Lenard's measurements, the longest effective wavelength for air is about $1.9 \times 10^{-5}$ cm, so that

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

We can also obtain an upper bound for the ionization energy in attenuated gases. According to J. Stark, the smallest measured ionization potential (for platinum anodes) is about 10 volts for air. We thus obtain $9.6 \times 10^{12}$ as the upper bound for $J$ which is very nearly the same as the value found above. There is another consequence, the experimental verification of which appears to me to be of great importance. If each quantum ionizes just one molecule, then the relationship $j = L/R\beta\nu$ must hold between the quantity of light $L$ and the number, $j$, of gram molecules ionized by this amount of light. If our picture is correct, this equation must hold for every gas which (for the given frequency) shows no appreciable absorption unaccompanied by ionization.

oooooooo

Max Born, in an essay honoring Einstein on his seventieth birthday, wrote that Vol. 17 of the *Annalen der Physik* of 1905 is "one of the most remarkable volumes in the whole scientific literature. It contains three papers by Einstein, each dealing with a different subject and each today acknowledged to be a masterpiece, the source of a new branch of physics. These three subjects, in order of pages, are: theory of photons, Brownian motion, and relativity." When these papers were published, Einstein was still a clerk in the Swiss patent office, and his only contact with the main stream of physics and the great physicists of that period was through some of their original papers and the standard treatises by such men as Mach, Kirchhoff, Helmholtz, Hertz, and others that were available at the time.

Considering the magnitude of Einstein's achievement in writing any one of the three papers mentioned by Born, we are amazed that this could have been done without direct contact with other physicists of that period. If one reads Einstein's *Autobiographical Notes*, excerpts of which are presented in this section, it becomes somewhat clearer as to just why he departed from the classical approach to space and time and how he arrived at his theory of relativity. Remaining outside the influence of the dominant physicists of the late nineteenth and early twentieth century, Einstein was free to speculate to his heart's desire and to wander along forbidden intellectual paths. Moreover, he was fortunate in beginning to probe at a time when serious disagreement existed between classical theory and experimental data and when the old mechanistic ideas (that all the phenomena of nature could be explained by means of Newtonian mechanics) were being challenged by the success of Maxwell's theory.

The work of Faraday, Maxwell, and Hertz had introduced into man's concept of the universe an entity that could exist in space quite independently of palpable matter, namely, the radiation field. As Einstein remarked in 1949,

> The factor which finally succeeded, after long hesitation, to bring the physicists slowly around to give up the faith in the possibility that all of physics could be founded upon Newton's mechanics, was the electrodynamics of Faraday and Maxwell. For this theory and its confirmation by Hertz's experiment showed that there are electromagnetic phenomena which by their very nature are detached from every ponderable matter—namely the waves in empty space which consist of electromagnetic fields. If mechanics was to be maintained as the foundation of physics, Maxwell's equations had to be interpreted mechanically. This was zealously but fruitlessly attempted, while the equations were proving themselves fruitful in mounting degree. One got used to operating with those fields as independent substances without finding it necessary to give one's self an account of their mechanical nature; thus mechanics as the basis of physics was being abandoned, almost unnoticeably, because its adaptability to the facts presented itself finally as hopeless.[2]

It was precisely in analyzing the incompatibility between electromagnetic phenomena, or light, and the classical laws of mechanics that Einstein realized that a fundamental change was necessary in our concepts of space and time. His primary interest in pursuing this analysis was to understand the manner in which light interacts with rapidly moving media. Put differently, how would an observer moving with a very great speed see radiative processes unfold themselves?

Most remarkable about Einstein's work was his willingness to give up the most cherished ideas of physicists, to set out boldly along new paths, and to reformulate problems in terms of the most elementary ideas. He states in his autobiographical notes that he realized shortly after 1900, i.e., after Planck's trail-blazing work, that . . .

> neither mechanics nor thermodynamics could (except in limiting cases) claim exact validity. By and by I despaired of the possibility of discovering the true laws by means of constructive efforts based on the known facts. The longer and the more despairingly I tried, the more I came to the conviction that only the discovery of a universal formal principle could lead us to assured results. The example I saw before me was thermodynamics. The general principle was there given in terms of the theorem: the laws of nature are such that it is impossible to construct a *perpetuum mobile* (of the first and second kind). How then could such a universal principle be found? After ten years of reflection such a principle resulted from a paradox upon which I had already hit at the age of sixteen.[3]

[2] Einstein, "Autobiographical Notes," in Paul Schilpp, *Albert Einstein, Philosopher-Scientist* ("The Library of Living Philosophers"; Evanston, Ill.: Library of Living Philosophers, 1949), Vol. 7, pp. 25, 27.



And here Einstein describes the first of a series of *Gedanken* (carried out in the mind of the scientist) experiments for which he is famous. These experiments, although extremely simple, were powerful tools in his hands and revealed to him the basic physical ideas involved in the phenomenon he was analyzing. The *Gedanken* experiment that occurred to Einstein at the age of sixteen dealt with the way in which a beam of light would appear to an observer traveling with the speed of light. He says,

> If I pursue a beam of light with the velocity [of light in a vacuum], I should observe such a beam of light as a spatially oscillatory electromagnetic field at rest. However, there seems to be no such thing, whether on the basis of experience or according to Maxwell's equations. From the very beginning it appeared to me intuitively clear that, judging from the standpoint of such an observer, everything would have to happen according to the same laws as for an observer who, relative to the earth, was at rest.[4]

This is the beginning of the principle of invariance (the laws of nature should appear the same to all observers moving with uniform speed with respect to each other) that Einstein was to make the basis of all his work in relativity. In this *Gedanken* experiment, Einstein had already formulated the idea that an observer could not travel with the speed of light, for if he did so, he would observe a stationary electromagnetic field.

A more profound analysis of the unique role of the speed of light in nature finally convinced him that this feature of light could be understood only if one revised one's concepts of space and time. He states it very clearly in the following paragraph:

> One sees that in this paradox [the way a beam of light would appear to an observer traveling with the speed of light] the germ of the special theory of relativity is already contained. Today everyone knows, of course, that all attempts to clarify this paradox satisfactorily were condemned to failure as long as the axiom of the absolute character of time, viz., of simultaneity, unrecognizedly was anchored in the unconscious. Clearly to recognize this axiom and its arbitrary character really implies already the solution of the problem. The type of critical reasoning which was required for the discovery

[3] *Ibid.*, p. 53.

[4] *Ibid.*, p. 53.

of this central point was decisively furthered, in my case, especially by the reading of David Hume's and Ernst Mach's philosophical writings.[5]

Einstein eventually realized that further progress could be made in resolving the above paradox only by first carefully analyzing such apparently simple ideas as the distance and the time interval between two events.

In classical physics, that is, before the advent of the special theory of relativity, two observers moving with uniform speed with respect to each other could compare their separate descriptions of events in the universe by means of simple mathematical formulas that connected their two coordinate systems (transformation of coördinates). Thus, any law of nature expressed mathematically by one observer in terms of his coordinate system could be translated into the mathematical language of the other observer by means of these transformations. Since all observers moving with uniform speed with respect to each other must be taken as equal in the eyes of nature, the transformation of a law from one coordinate system to another should leave the law unaltered (principle of relativity or the principle of invariance); yet these transformation equations do not leave an experimentally verified law, namely, the constancy of the speed of light (based upon the Michelson-Morley experiment and Einstein's *Gedanken* experiment) unaltered when one passes from one system to another. In other words, the constancy of the speed of light for all observers, regardless of the speed with which they are moving with respect to the earth, is in direct contradiction to the transformation equations of classical physics.

Einstein became aware of this and realized that if one accepted the constancy of the speed of light as an experimentally and, in terms of Maxwell's equations, a theoretically established law of nature, then one would have to replace the classical transformation equations by "relations of a new type (Lorentz transformation)." Einstein saw further that any set of relations (that is, transformation equations) that enables one to pass from a description of the laws of nature and events in one coordinate system to a description in another coordinate system involves a very definite concept of space and time; it has meaning only in terms of the specific way in which the measurements of distance and time are introduced.

The Newtonian way of looking at space and time, as absolute and independent entities in our universe, was based upon a very definite way of interpreting measuring rods and clocks. As long as one insisted on the correctness of this classical picture, it was impossible to fit the constancy of the speed of light into the framework in which the laws of nature are taken as independent of the choice of the inertial frame of reference. The only way these two ideas can be made compatible is by introducing a new

[5] *Ibid.*, p. 53.



hypothesis concerning "the actual behavior of moving measuring-rods and clocks, which can be experimentally validated or disproved." Einstein was led to a reevaluation of the role of measuring rods and clocks in establishing the laws of physics and of how our knowledge of their behavior in motion would be affected if we accept the constancy of the speed of light.

In classical physics it was taken as an a priori truth that if different observers in the universe were to order the events of history along a time axis and were to specify their positions and the distances between them, then this would always be the same arrangement regardless of how the observers were moving with respect to each other. According to this picture space and time are pictured as entirely independent of each other. Einstein's great contribution was to demonstrate that the constancy of the speed of light brought with it a new picture of space and time in which the two are fused into a single continuum with the space and time parts having different aspects for different observers.

By combining the space and the time intervals between two events into a single space-time interval, Einstein showed how the classical transformation equations would give way to the Lorentz transformations. All these developments were presented in a small book by Einstein written in 1916 entitled *Relativity, the Special and the General Theory*. Excerpts taken from this book that present the salient features of the development of special relativity are reproduced later in this chapter. The transformation equations Einstein obtained (they are the basis of the special theory of relativity) from this general analysis of space and time are identical with those Lorentz obtained by analyzing the behavior of his theoretical electrons. But the Lorentz derivation did not lead to a revision of the concepts of space and time, since Lorentz always considered his equations applicable only to electrons; indeed, from the way Lorentz derived these equations there was no justification for drawing more comprehensive conclusions from them. Lorentz, in fact, long persisted in the idea that his transformation equations were a peculiarity of electronic behavior and that the classical transformations were valid in the general case.

## Autobiographical Notes [6]

. . . The most fascinating subject at the time that I was a student was Maxwell's theory. What made this theory appear revolutionary was the transition from forces at a distance to fields as fundamental variables.

[6] *Ibid.*, pp. 33–53.

The incorporation of optics into the theory of electromagnetism, with its relation of the speed of light to the electric and magnetic absolute system of units as well as the relation of the refraction coëfficient to the dielectric constant, the qualitative relation between the reflection coëfficient and the metallic conductivity of the body—it was like a revelation. Aside from the transition to field-theory, i.e., the expression of the elementary laws through differential equations, Maxwell needed only one single hypothetical step—the introduction of the electrical displacement current in the vacuum and in the dielectrica and its magnetic effect, an innovation which was almost prescribed by the formal properties of the differential equations. In this connection I cannot suppress the remark that the pair Faraday-Maxwell has a most remarkable inner similarity with the pair Galileo-Newton—the former of each pair grasping the relations intuitively, and the second one formulating those relations exactly and applying them quantitatively.

What rendered the insight into the essence of electromagnetic theory so much more difficult at that time was the following peculiar situation. Electric or magnetic "field intensities" and "displacements" were treated as equally elementary variables, empty space as a special instance of a dielectric body. *Matter* appeared as the bearer of the field, not *space*. By this it was implied that the carrier of the field could have velocity, and this was naturally to apply to the "vacuum" (ether) also. Hertz's electrodynamics of moving bodies rests entirely upon this fundamental attitude.

It was the great merit of H. A. Lorentz that he brought about a change here in a convincing fashion. In principle a field exists, according to him, only in empty space. Matter—considered as atoms—is the only seat of electric charges; between the material particles there is empty space, the seat of the electromagnetic field, which is created by the position and velocity of the point charges which are located on the material particles. Dielectricity, conductivity, etc., are determined exclusively by the type of mechanical tie connecting the particles, of which the bodies consist. The particle-charges create the field, which, on the other hand, exerts forces upon the charges of the particles, thus determining the motion of the latter according to Newton's law of motion. If one compares this with Newton's system, the change consists in this: action at a distance is replaced by the field, which thus also describes the radiation. Gravitation is usually not taken into account because of its relative smallness; its consideration, however, was always possible by means of the enrichment of the structure of the field, i.e., expansion of Maxwell's law of the field. The physicist of the present generation regards the point of view achieved by Lorentz as the only possible one; at that time, however, it was a surprising and audacious step, without which the later development would not have been possible.



If one views this phase of the development of theory critically, one is struck by the dualism which lies in the fact that the material point in Newton's sense and the field as continuum are used as elementary concepts side by side. Kinetic energy and field-energy appear as essentially different things. This appears all the more unsatisfactory inasmuch as, according to Maxwell's theory, the magnetic field of a moving electric charge represents inertia. Why not then *total* inertia? Then only field-energy would be left, and *the particle would be merely an area of special density of field-energy.** In that case one could hope to deduce the concept of the mass-point together with the equations of the motion of the particles from the field equations—the disturbing dualism would have been removed.

H. A. Lorentz knew this very well. However, Maxwell's equations did not permit the derivations of the equilibrium of the electricity which constitutes a particle. Only other, nonlinear field equations could possibly accomplish such a thing. But no method existed by which this kind of field equations could be discovered without deteriorating into adventurous arbitrariness. In any case one could believe that it would be possible by and by to find a new and secure foundation for all of physics upon the path which had been so successfully begun by Faraday and Maxwell.

Accordingly, the revolution begun by the introduction of the field was by no means finished. Then it happened that, around the turn of the century, independently of what we have just been discussing, a second fundamental crisis set in, the seriousness of which was suddenly recognized due to Max Planck's investigations into heat radiation (1900). The history of this event is all the more remarkable because, at least in its first phase, it was not in any way influenced by any surprising discoveries of an experimental nature.

On thermodynamic grounds Kirchhoff had concluded that the energy density and the spectral composition of radiation in a *Hohlraum*, surrounded by impenetrable walls of the temperature $T$, would be independent of the nature of the walls. That is to say, the nonchromatic density of radiation $\rho$ is a universal function of the frequency $\nu$ and of the absolute temperature $T$. Thus arose the interesting problem of determining this function $\rho(\nu,T)$. What could theoretically be ascertained about this function? According to Maxwell's theory the radiation had to exert a pressure on the walls, determined by the total energy density. From this Boltzmann concluded, by means of pure thermodynamics, that the entire energy density of the radiation ($\int \rho d\nu$) is proportional to $T^4$. In this way he found a theoretical justification of a law which had previously been discovered empirically by Stefan, i.e., in this way he connected this empirical law with the basis of Maxwell's theory. Thereafter, by way of an ingenious thermo-

* Editors italics.

dynamic consideration, which also made use of Maxwell's theory, W. Wien found that the universal function $\rho$ of the two variables $\nu$ and $T$ would have to be of the form

$$\rho \approx \nu^3 f\left(\frac{\nu}{T}\right),$$

whereby $f(\nu/T)$ is a universal function of one variable $\nu/T$ only. It was clear that the theoretical determination of this universal function $f$ was of fundamental importance—this was precisely the task which confronted Planck. Careful measurements had led to a very precise empirical determination of the function $f$. Relying on those empirical measurements, he succeeded in the first place in finding a statement which rendered the measurements very well indeed:

$$\rho = \frac{8\pi h\nu^3}{c^3} \frac{1}{\exp(h\nu/kT) - 1},$$

whereby $h$ and $k$ are two universal constants, the first of which led to quantum theory. Because of the denominator this formula looks a bit queer. Was it possible to derive it theoretically? Planck actually did find a derivation, the imperfections of which remained at first hidden, which latter fact was most fortunate for the development of physics. If this formula was correct, it permitted, with the aid of Maxwell's theory, the calculation of the average energy $E$ of a quasi-monochromatic oscillator within the field of radiation:

$$E = \frac{h\nu}{\exp(h\nu/kT) - 1}.$$

Planck preferred to attempt calculating this latter magnitude theoretically. In this effort, thermodynamics, for the time being, proved no longer helpful, and neither did Maxwell's theory. The following circumstance was unusually encouraging in this formula. For high temperatures (with a fixed $\nu$) it yielded the expression

$$E = kT.$$

This is the same expression as the kinetic theory of gases yields for the average energy of a mass-point which is capable of oscillating elastically in one dimension. For in kinetic gas theory one gets

$$E = (R/N)T,$$



whereby $R$ means the constant of the equation of state of a gas and $N$ the number of molecules per mol, from which constant one can compute the absolute size of the atom. Putting these two expressions equal to each other one gets

$$N = R/k.$$

The one constant of Planck's formula consequently furnishes exactly the correct size of the atom. The numerical value agreed satisfactorily with the determinations of $N$ by means of kinetic gas theory, even though these latter were not very accurate.

This was a great success, which Planck clearly recognized. But the matter has a serious drawback, which Planck fortunately overlooked at first. For the same considerations demand in fact that the relation $E = kT$ would also have to be valid for low temperatures. In that case, however, it would be all over with Planck's formula and with the constant $h$. From the existing theory, therefore, the correct conclusion would have been: the average kinetic energy of the oscillator is either given incorrectly by the theory of gases, which would imply a refutation of [statistical] mechanics; or else the average energy of the oscillator follows incorrectly from Maxwell's theory, which would imply a refutation of the latter. Under such circumstances it is most probable that both theories are correct only at the limits, but are otherwise false; this is indeed the situation, as we shall see in what follows. If Planck had drawn this conclusion, he probably would not have made his great discovery, because the foundation would have been withdrawn from pure deductive reasoning.

Now back to Planck's reasoning. On the basis of the kinetic theory of gases Boltzmann had discovered that, aside from a constant factor, entropy is equivalent to the logarithm of the "probability" of the state under consideration. Through this insight he recognized the nature of courses of events which, in the sense of thermodynamics, are "irreversible." Seen from the molecular-mechanical point of view, however, all courses of events are reversible. If one calls a molecular-theoretically defined state a microscopically described one, or, more briefly, micro-state, and a state described in terms of thermodynamics a macro-state, then an immensely large number (Z) of states belong to a macroscopic condition. Z then is a measure of the probability of a chosen macro-state. This idea appears to be of outstanding importance also because of the fact that its usefulness is not limited to microscopic description on the basis of mechanics. Planck recognized this and applied the Boltzmann principle to a system which consists of very many resonators of the same frequency $\nu$. The macroscopic situation is given through the total energy of the oscillation of all resonators, a micro-condition through determination of the (instantaneous) energy of each individual resonator. In order then to be able to express the number of the micro-states belonging to a macro-state by means of a finite number, he [Planck] divided the total energy into a large but finite number of identical energy-elements $\epsilon$ and asked: in how many ways can these energy-elements be divided among the resonators. The logarithm of this number, then, furnishes the entropy and thus (via thermodynamics) the temperature of the system. Planck got his radiation-formula if he chose his energy-elements $\epsilon$ of the magnitude $\epsilon = h\nu$. The decisive element in doing this lies in the fact that the result depends on taking for $\epsilon$ a definite finite value, i.e., that one does not go to the limit $\epsilon = 0$. This form of reasoning does not make obvious the fact that it contradicts the mechanical and electrodynamic basis, upon which the derivation otherwise depends. Actually, however, the derivation presupposes implicitly that energy can be absorbed and emitted by the individual resonator only in "quanta" of magnitude $h\nu$, i.e., that the energy of a mechanical structure capable of oscillations as well as the energy of radiation can be transferred only in such quanta—in contradiction to the laws of mechanics and electrodynamics. The contradiction with dynamics was here fundamental; whereas the contradiction with electrodynamics could be less fundamental. For the expression for the density of radiation-energy, although it is *compatible* with Maxwell's equations, is not a necessary consequence of these equations. That this expression furnishes important average-values is shown by the fact that the Stefan-Boltzmann law and Wien's law, which are based on it, are in agreement with experience.

All of this was quite clear to me shortly after the appearance of Planck's fundamental work; so that, without having a substitute for classical mechanics, I could nevertheless see to what kind of consequences this law of temperature-radiation leads for the photo-electric effect and for other related phenomena of the transformation of radiation-energy, as well as for the specific heat of (especially) solid bodies. All my attempts, however, to adapt the theoretical foundation of physics to this [new type of] knowledge failed completely. It was as if the ground had been pulled out from under one, with no firm foundation to be seen anywhere, upon which one could have built. That this insecure and contradictory foundation was sufficient to enable a man of Bohr's unique instinct and tact to discover the major laws of the spectral lines and of the electron-shells of the atoms together with their significance for chemistry appeared to me like a miracle—and appears to me as a miracle even today. This is the highest form of musicality in the sphere of thought.

My own interest in those years was less concerned with the detailed consequences of Planck's results, however important these might be. My major question was: What general conclusions can be drawn from the radiation-formula concerning the structure of radiation and even more generally concerning the electro-magnetic foundation of physics? Before I take this up, I must briefly mention a number of investigations which relate to the Brownian motion and related objects (fluctuation-phenomena) and which in essence rest upon classical molecular mechanics. Not acquainted with the earlier investigations of Boltzmann and Gibbs, which had appeared earlier and actually exhausted the subject, I developed the statistical mechanics and the molecular-kinetic theory of thermodynamics which was based on the former. My major aim in this was to find facts which would guarantee as much as possible the existence of atoms of definite finite size. In the midst of this I discovered that, according to atomist theory, there would have to be a movement of suspended microscopic particles open to observation, without knowing that observations concerning the Brownian motion were already long familiar. The simplest derivation rested upon the following consideration. If the molecular-kinetic theory is essentially correct, a suspension of visible particles must possess the same kind of osmotic pressure fulfilling the laws of gases as a solution of molecules. This osmotic pressure depends upon the actual magnitude of the molecules, i.e., upon the number of molecules in a gram-equivalent. If the density of the suspension is inhomogeneous, the osmotic pressure is inhomogeneous, too, and gives rise to a compensating diffusion, which can be calculated from the well-known mobility of the particles. This diffusion can, on the other hand, also be considered as the result of the random displacement—unknown in magnitude originally—of the suspended particles due to thermal agitation. By comparing the amounts obtained for the diffusion current from both types of reasoning one reaches quantitatively the statistical law for those displacements, i.e., the law of the Brownian motion. The agreement of these considerations with experience together with Planck's determination of the true molecular size from the law of radiation (for high temperatures) convinced the sceptics, who were quite numerous at that time (Ostwald, Mach) of the reality of atoms. The antipathy of these scholars towards atomic theory can indubitably be traced back to their positivistic philosophical attitude. This is an interesting example of the fact that even scholars of audacious spirit and fine instinct can be obstructed in the interpretation of facts by philosophical prejudices. The prejudice—which has by no means died out in the meantime—consists in the faith that facts by themselves can and should yield scientific knowledge without free conceptual construction. Such a misconception is possible only because one does not easily become aware of the free choice of such concepts, which, through verification and long usage, appear to be immediately connected with the empirical material.

The success of the theory of the Brownian motion showed again conclusively that classical mechanics always offered trustworthy results whenever it was applied to motions in which the higher time derivatives of velocity are negligibly small. Upon this recognition a relatively direct method can be based which permits us to learn something concerning the constitution of radiation from Planck's formula. One may conclude in fact that, in a space filled with radiation, a (vertically to its plane) freely moving, quasi-monochromatically reflecting mirror would have to go through a kind of Brownian movement, the average kinetic energy of which equals $\frac{1}{2}(R/N)T$ ($R$ = constant of the gas-equation for one gram-molecule, $N$ equals the number of the molecules per mol, $T$ = absolute temperature). If radiation were not subject to local fluctuations, the mirror would gradually come to rest, because, due to its motion, it reflects more radiation on its front than on its reverse side. However, the mirror must experience certain random fluctuations of the pressure exerted upon it due to the fact that the wave-packets, constituting the radiation, interfere with one another. These can be computed from Maxwell's theory. This calculation, then, shows that these pressure variations (especially in the case of small radiation-densities) are by no means sufficient to impart to the mirror the average kinetic energy $\frac{1}{2}(R/N)T$. In order to get this result one has to assume rather that there exists a second type of pressure variations, which can not be derived from Maxwell's theory, which corresponds to the assumption that radiation energy consists of indivisible point-like localized quanta of the energy $h\nu$ (and of momentum ($h\nu/c$), ($c$ = velocity of light)), which are reflected undivided. This way of looking at the problem showed in a drastic and direct way that a type of immediate reality has to be ascribed to Planck's quanta, that radiation must, therefore, possess a kind of molecular structure in energy, which of course contradicts Maxwell's theory. Considerations concerning radiation which are based directly on Boltzmann's entropy-probability-relation (probability taken equal to statistical temporal frequency) also lead to the same result. This double nature of radiation (and of material corpuscles) is a major property of reality, which has been interpreted by quantum-mechanics in an ingenious and amazingly successful fashion. This interpretation, which is looked upon as essentially final by almost all contemporary physicists, appears to me as only a temporary way out; a few remarks to this [point] will follow later.

Reflections of this type made it clear to me as long ago as shortly after 1900, i.e., shortly after Planck's trail-blazing work, that neither mechanics nor thermodynamics could (except in limiting cases) claim exact validity. By and by I despaired of the possibility of discovering the true laws by means of constructive efforts based on known facts. The longer and the more despairingly I tried, the more I came to the conviction that only the discovery of a universal formal principle could lead us to assured results. The example I saw before me was thermodynamics. The general principle was there given in the theorem: the laws of nature are such that it is impossible to construct a *perpetuum mobile* (of the first and second kind). How, then, could such a universal principle be found? After ten years of reflection such a principle resulted from a paradox upon which I had already hit at the age of sixteen: If I pursue a beam of light with the velocity $c$ (velocity of light in a vacuum), I should observe such a beam of light as a spatially oscillatory electromagnetic field at rest. However, there seems to be no such thing, whether on the basis of experience or according to Maxwell's equations. From the very beginning it appeared to me intuitively clear that, judged from the standpoint of such an observer, everything would have to happen according to the same laws as for an observer who, relative to the earth, was at rest. For how, otherwise, should the first observer know, i.e., be able to determine that he is in a state of fast uniform motion?



One sees that in this paradox the germ of the special relativity theory is already contained. Today everyone knows, of course, that all attempts to clarify this paradox satisfactorily were condemned to failure as long as the axiom of the absolute character of time, viz., of simultaneity, unrecognizedly was anchored in the unconscious. Clearly to recognize this axiom and its arbitrary character really implies already the solution of the problem. . . .

## The Special Theory of Relativity [7]

### SPACE AND TIME IN CLASSICAL MECHANICS

THE PURPOSE OF MECHANICS IS to describe how bodies change their position in space with "time." I should load my conscience with grave sins against the sacred spirit of lucidity were I to formulate the aims of mechanics in this way, without serious reflection and detailed explanations. Let us proceed to disclose these sins.

It is not clear what is to be understood here by "position" and "space." I stand at the window of a railway carriage which is travelling uniformly, and drop a stone on the embankment, without throwing it. Then, disregarding the influence of the air resistance, I see the stone descend in a straight line. A pedestrian who observes the misdeed from the footpath notices that the stone falls to earth in a parabolic curve. I now ask: Do the "positions" traversed by the stone lie "in reality" on a straight line or on a parabola? Moreover, what is meant here by motion "in space"? . . . The

[7] Einstein, *Relativity: The Special and General Theory*, trans. R. W. Lawson (New York: Henry Holt, 1920), pp. 9–56.