

Zero-point energy in early quantum theory

P. W. Milonni, and M. -L. Shih

Citation: *American Journal of Physics* **59**, 684 (1991); doi: 10.1119/1.16772

View online: <http://dx.doi.org/10.1119/1.16772>

View Table of Contents: <http://aapt.scitation.org/toc/ajp/59/8>

Published by the [American Association of Physics Teachers](#)

Articles you may be interested in

[Any classical description of nature requires classical electromagnetic zero-point radiation](#)
American Journal of Physics **79**, 1163 (2011); 10.1119/1.3630939

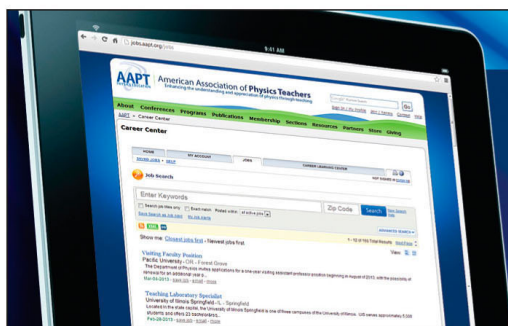
[Essentials of the Casimir effect and its computation](#)
American Journal of Physics **59**, 711 (1998); 10.1119/1.16749

[The Quantum Vacuum: An Introduction to Quantum Electrodynamics](#)
American Journal of Physics **62**, 1154 (1998); 10.1119/1.17618

[Casimir effect from a scattering approach](#)
American Journal of Physics **83**, 156 (2015); 10.1119/1.4896197

QUANTUM MEASUREMENTS
American Journal of Physics **85**, 5 (2016); 10.1119/1.4967925

[The Casimir effect from a condensed matter perspective](#)
American Journal of Physics **77**, 1055 (2009); 10.1119/1.3194050



American Association of **Physics Teachers**

Explore the **AAPT Career Center** –
access hundreds of physics education and
other STEM teaching jobs at two-year and
four-year colleges and universities.

<http://jobs.aapt.org>



Zero-point energy in early quantum theory^{a)}

P. W. Milonni and M. -L. Shih

Theoretical Division, Mail Stop B-212, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

(Received 20 June 1990; accepted for publication 21 February 1991)

In modern physics the vacuum is not a tranquil void but a quantum state with fluctuations having observable consequences. The present concept of the vacuum has its roots in the zero-point energy of harmonic oscillators and the electromagnetic field, and arose before the development of the formalism of quantum mechanics. This article discusses these roots in the blackbody research of Planck and Einstein in 1912–1913, and the relation to Bose–Einstein statistics and the first indication of wave–particle duality uncovered by Einstein’s fluctuation formula. Also considered are the Einstein–Stern theory of specific heats, which invoked zero-point energy in a way which turned out to be incorrect, and the experimental implications of zero-point energy recognized by Mulliken and Debye in vibrational spectroscopy and x-ray diffraction.

“The existence of a zero-point energy of size $\frac{1}{2} h\nu$ [is] probable.”—Albert Einstein and Otto Stern (1913).

I. INTRODUCTION

According to contemporary physics, the Universe is made up of matter fields, whose quanta are fermions, and force fields, whose quanta are bosons. All these fields have vacuum fluctuations and zero-point energy. Although the concept of zero-point energy arises in introductory courses on quantum mechanics, usually in connection with the harmonic oscillator, it is typically brushed aside as one of the formal peculiarities of quantum theory, and even in advanced treatises there is seldom any discussion of its theoretical significance or experimental relevance.

The concept of zero-point energy actually arose well before the development of the quantum formalism in 1925–1926; quantum mechanics confirmed that zero-point energy is physically real. Zero-point energy was especially in vogue during the second decade of this century, probably in part because of the great interest at the time in low-temperature phenomena. As far as we have been able to determine, zero-point energy first appeared in Planck’s “second theory” of blackbody radiation.¹ It was quickly adopted by Einstein and others.

This article is a discussion of these roots of the concept of zero-point energy. We do not proffer any sort of rigorous historical analysis, but only a glimpse into some of the early physics of energy at the absolute zero of temperature.

We begin in Sec. II with some background on the blackbody problem, including the failure of the classical model of an atom as an elastically bound electron. In Sec. III we review Planck’s “first theory” in a way that leads us conveniently, in Sec. IV, to Planck’s later introduction of zero-point energy *for the elastically bound electron but not the electromagnetic field*. The material in Secs. II and III is readily available in various historical analyses,² but we have found concise discussions of Planck’s original arguments to be surprisingly rare in the pedagogical literature. For this reason and because of its obvious historical importance and relevance to our subject, we feel justified in including this material here. In Secs. V and VI we discuss some work of Einstein where again zero-point energy was assumed only for material oscillators. We speculate as to why the concept of zero-point energy was not extended to the electromagnetic field.

We show in Sec. VII that the well-known “particle term” in the Einstein fluctuation formula may be regarded

as a consequence of zero-point energy. We elaborate on this point in Secs. VIII and IX, and discuss the connection to Bose–Einstein statistics and spontaneous emission. In Sec. X we describe the Einstein–Stern theory of specific heats, where zero-point energy plays a crucial role in obtaining a specific heat with the correct low-temperature behavior. We explain why this ingenious theory turned out to be incorrect. Sections XI and XII focus on the role of zero-point energy in x-ray diffraction and molecular vibrations, as first explained by Debye and Mulliken, respectively. Section XIII is a brief summary with some additional notes about the history of zero-point energy.

Although much of the material herein is perhaps not suitable for inclusion in standard courses, we believe that parts of it could usefully supplement the standard texts and, furthermore, that it might be appropriate for seminars or research projects for theoretically inclined students. But our main purpose in this article is to trace the origins of zero-point energy in a unified way that, to our knowledge, is not available elsewhere in the literature.

II. THE BLACKBODY PROBLEM AND THE CLASSICAL OSCILLATOR MODEL OF AN ATOM

In 1860 Kirchoff derived a law relating the radiative and absorptive strengths of a body held at a fixed temperature T .² According to Kirchoff’s law, the ratio of the radiative intensity to the absorption coefficient for radiation of wavelength λ is the same for all bodies at temperature T and defines a universal function $F(\lambda, T)$. This led to the abstraction of an ideal *blackbody* for which the absorption coefficient is unity at all wavelengths, corresponding to total absorption. Thus, $F(\lambda, T)$ characterizes the radiative strength at wavelength λ of a blackbody at temperature T . The problem was to determine the universal function $F(\lambda, T)$.

An important step was taken in 1884 by Boltzmann, who invoked several aspects of Maxwell’s electromagnetic theory. The most important of these for the present discussion is the result that isotropic radiation exerts on a perfectly reflecting surface a pressure $u/3$, where u is the energy density of the radiation.³ Boltzmann considered blackbody radiation confined in a cylinder of volume V , one end of which is a perfectly reflecting piston. The radiation pressure on the piston increases the volume by dV , and in order to maintain a constant temperature, an amount of heat,

$$dQ = dU + P dV = d(uV) + \frac{1}{3}u dV = V du + \frac{4}{3}u dV, \quad (1)$$

must be added, according to the first law of thermodynamics. Kirchoff's law implies that the total energy density u over all wavelengths is a function only of T , so that

$$dQ = V \frac{du}{dT} dT + \frac{4}{3} u dV. \quad (2)$$

Associated with the expansion of the cylinder is an increase in entropy by

$$dS = \frac{1}{T} dQ = \frac{V}{T} \frac{du}{dT} dT + \frac{4}{3} \frac{u}{T} dV, \quad (3)$$

which, according to the second law of thermodynamics, is an exact differential. Thus,

$$\frac{\partial S}{\partial T} = \frac{V}{T} \frac{du}{dT}, \quad \frac{\partial S}{\partial V} = \frac{4}{3} \frac{u}{T} \quad (4)$$

and

$$\frac{\partial^2 S}{\partial T \partial V} = \frac{\partial^2 S}{\partial V \partial T} = \frac{1}{T} \frac{du}{dT} = \frac{4}{3} \frac{d}{dT} \left(\frac{u}{T} \right), \quad (5)$$

from which it follows that $du/dT = 4u/T$ and

$$u = bT^4 \quad (\text{Stefan-Boltzmann law}), \quad (6)$$

where b is a universal constant. Stefan in 1879 had in fact suggested such a relation from an analysis of experimental data.

The Stefan-Boltzmann law stands in conflict with elementary classical models of equilibrium between radiation and matter. Consider the classical oscillator model of an atom, where an electron is assumed to be bound by an elastic restoring force.⁴ If $\rho(\nu)d\nu$ denotes the energy per unit volume of radiation in the frequency interval $[\nu, \nu + d\nu]$, then the rate at which the atom absorbs energy from the radiation field may be shown to be given by the formula (see Appendix A)

$$\dot{W}_A = (\pi e^2/3m)\rho(\nu_0), \quad (7)$$

where \dot{W}_A is the electron energy, e and m are its charge and mass, respectively, and ν_0 is the natural oscillation frequency of the electron in the atom. The rate at which the electron radiates electromagnetic energy \dot{W}_{EM} is given by the well-known classical Larmor formula:

$$\dot{W}_{EM} = 2e^2 a^2/3c^3, \quad (8)$$

where a is the acceleration of the electron. For oscillation at frequency $\nu_0 = \omega_0/2\pi$, $a = -\omega_0^2 x$ and

$$\dot{W}_{EM} = (32\pi^4 e^2 \nu_0^4/3c^3)x^2, \quad (9)$$

where x is the electron displacement from its equilibrium position in the classical oscillator model of the atom. Now, according to the virial theorem of classical mechanics, the average potential energy $\frac{1}{2}m\omega_0^2 x^2$ of the (one-dimensional) electron oscillator is equal to the average kinetic energy, and their sum is the total oscillator energy U . In a state of equilibrium between radiation and matter, furthermore, the energy absorption rate (7) should equal the emission rate (9). Thus,

$$\rho(\nu_0) = (8\pi\nu_0^2/c^3)(m\omega_0^2 x^2) = (8\pi\nu_0^2/c^3)U, \quad (10)$$

or, more generally,

$$\rho(\nu) = (8\pi\nu^2/c^3)U, \quad (11)$$

for a blackbody, which absorbs at all frequencies ν . Finally, the equipartition theorem of classical statistical mechanics demands that the average value of U in thermal equilibrium is kT , where k is Boltzmann's constant, so that the

spectral energy density of thermal radiation must be

$$\rho(\nu) = (8\pi\nu^2/c^3)kT \quad (\text{Rayleigh-Jeans distribution}). \quad (12)$$

The total electromagnetic energy density

$$u = \int_0^\infty \rho(\nu) d\nu \quad (13)$$

violates the Stefan-Boltzmann law. Furthermore, the Rayleigh-Jeans law suffers from the ultraviolet catastrophe: u diverges when (12) is used for $\rho(\nu)$. Equation (11) was derived by Planck and played a very important role in his work on the blackbody problem.

Equation (12) was first deduced in a less explicit form by Rayleigh in 1900.⁵ Although the derivation just outlined might be criticized for its reliance on a particular model of an atomic electron, it is easy to derive the Rayleigh-Jeans distribution on more general classical grounds. An electromagnetic field mode of frequency ν is basically just a linear harmonic oscillator which, according to the classical equipartition theorem, has an average energy kT at thermal equilibrium.⁶ Since the number of modes per unit volume in the frequency interval $[\nu, \nu + d\nu]$ is $(8\pi\nu^2/c^3)d\nu$, the electromagnetic energy per unit volume in this frequency interval should be $(8\pi\nu^2/c^3)(kT)d\nu = \rho(\nu)d\nu$, which is the Rayleigh-Jeans law, independently of any particular model for the atoms with which the radiation is in thermal equilibrium. From this perspective the failure of classical theory, according to Kelvin and Rayleigh, must lie in its equipartition theorem.

Another classical result, due to Wien in 1893, must be mentioned. Wien basically followed Boltzmann's model of radiation contained in a cylinder with a piston, but included the Doppler shift of radiation reflected by the moving piston. This allowed radiant energy to be exchanged among different frequencies. Wien showed that the spectral energy density must follow the general form

$$\rho(\nu) = \nu^3 \phi_1(\nu/T) \quad (\text{Wien displacement law}), \quad (14)$$

or, in terms of wavelength,

$$\rho(\lambda) = \rho(\nu) \left| \frac{d\nu}{d\lambda} \right| = \lambda^{-5} \phi_2(\lambda T) \quad (\text{Wien displacement law}), \quad (15)$$

where ϕ_1 and ϕ_2 are undetermined functions. The Rayleigh-Jeans distribution obviously obeys Wien's "displacement law" (14).

A few years later, Wien presented arguments in support of the distribution

$$\rho(\lambda) = \alpha \lambda^{-5} e^{-\beta/\lambda T} \quad (\text{Wien distribution}), \quad (16)$$

where α and β are constants. A similar distribution function, with the factor λ^{-5} replaced by $\lambda^{-\gamma}$, had just been proposed by Paschen as a fit to his experimental data. Paschen's data indicated that γ was between 5 and 6, thus providing some support for the displacement law. Further measurements showed that γ was indeed close to 5.

Wien's arguments for (16) seem to have been guided more by the desired result than by physics. To wit, he made the peculiar assumption that the wavelength and intensity of the radiation from a given atom (or molecule) are determined only by that atom's velocity. This allowed him to adduce the exponential term in (16) from the factor $\exp(-mv^2/2kT)$ in the Maxwell-Boltzmann velocity distribution function. In any case, the Wien distribution

was soon to find a more secure provenance in Planck's work.

III. PLANCK'S FIRST THEORY

Given that Planck was an expert in thermodynamics, it is not surprising that his work on the blackbody problem emphasized the concept of entropy. In a series of papers in the late 1890s, Planck produced a derivation of the Wien distribution from general thermodynamical considerations plus an *assumption* that the entropy of a collection of radiators depends only on their total energy. An important result was the following relation between the entropy S and average energy U of an elementary radiator (or "molecule" for our purposes) in thermal equilibrium with radiation at temperature T :

$$\frac{\partial^2 S}{\partial U^2} = -\frac{A}{U}, \quad (17)$$

where for a given radiator A is a constant. From this equation and the general relation $\partial S/\partial U = 1/T$, it follows that

$$U = Be^{-1/AT}, \quad (18)$$

where B is another constant which, like A , may depend on the frequency of a given radiator. This result, together with (11), yields the radiation spectral energy density

$$\rho(\nu) = f(\nu)e^{-1/AT}, \quad (19)$$

where $f(\nu)$ is some function of ν . Wien's displacement law implies that $f(\nu)$ and A are proportional to ν^3 and ν^{-1} , respectively, so that

$$\rho(\nu) = C\nu^3 e^{-D\nu/T} \quad (C, D \text{ constants}) \quad (20)$$

or

$$\rho(\lambda) = \alpha\lambda^{-5} e^{-\beta/\lambda T} \quad (\alpha, \beta \text{ constants}), \quad (21)$$

which is the Wien distribution.

The Wien distribution, however, was soon found to be incorrect as experimentalists extended their spectral measurements to higher wavelengths. This was accomplished by the "residual rays" method, whereby longer wavelengths were isolated by multiple reflections off an appropriate crystal. In February 1900, Lummer and Pringsheim reported data that deviated from the Wien distribution by 40%–50% for wavelengths between 12 and 18 μm , and in October similar conclusions were reported by Rubens and Kurlbaum.

It was the work of his friend Rubens that led Planck to his formula for the spectral energy density of thermal radiation.⁷ In particular, the data indicated that $\rho(\nu)$ was proportional to the temperature T for small ν and large T . Planck found a formula with that behavior at small ν and which approximated the Wien distribution for large ν .

In a paper delivered at a meeting on 19 October, Planck presented his formula and provided some justification for it.^{2,7} For small ν and large T , the experimental result $\rho(\nu) \propto T$ and Eq. (11) imply $U \propto T$, and therefore, since $\partial S/\partial U = T^{-1}$, $\partial^2 S/\partial U^2 \propto U^{-2}$ and $S \propto \log U$. On the other hand, (17) leads to the Wien distribution, which has the correct form for large ν and small T . Planck therefore proposed the interpolation

$$\frac{\partial^2 S}{\partial U^2} = \frac{-A}{U(B+U)} \quad (A, B \text{ constants}). \quad (22)$$

According to Planck, Eq. (22) "is the simplest by far of all the expressions which yield S as a logarithmic function of U

(a condition which probability theory suggests) and which besides coincides with the Wien law for small values of U ." Using again the relation $\partial S/\partial U = 1/T$, Eq. (11), and the Wien displacement law, one obtains from (22) the spectral energy density

$$\rho(\lambda) = \alpha\lambda^{-5}/(e^{\beta/\lambda T} - 1) \quad (\alpha, \beta \text{ constants}). \quad (23)$$

This formula was found to agree with all the existing data. In order to give it "a real physical meaning," Planck began what he later described as "a few weeks of the most strenuous work of my life." The culmination of that work was the birth of quantum theory.

Planck's reasoning may be summarized as follows. Consider N radiators of frequency ν and total energy $U_N = NU = P\epsilon$, where P is a large integer and ϵ is some finite element of energy. The entropy $S_N = NS = k \log W_N$, where W_N is the number of ways in which the P energy elements can be distributed among the N radiators. If $N = P = 2$, for instance, then the different partitions of the energy between the two radiators are $(2\epsilon, 0)$, (ϵ, ϵ) , and $(0, 2\epsilon)$ if the energy elements are assumed to be indistinguishable. Under this assumption we have, in general,

$$W_N = \frac{(N-1+P)!}{P!(N-1)!}, \quad (24)$$

which is the number of ways in which P indistinguishable balls can be put into N distinguishable boxes. Stirling's approximation ($\log M! \cong M \log M - M$ for large M) then gives, for $N, P \gg 1$,

$$\begin{aligned} S &= \frac{k}{N} \log \frac{(N-1+P)!}{P!(N-1)!} \\ &\cong k \left[\left(1 + \frac{P}{N}\right) \log \left(1 + \frac{P}{N}\right) - \frac{P}{N} \log \frac{P}{N} \right] \\ &= k \left[\left(1 + \frac{U}{\epsilon}\right) \log \left(1 + \frac{U}{\epsilon}\right) - \frac{U}{\epsilon} \log \frac{U}{\epsilon} \right]. \end{aligned} \quad (25)$$

Thus

$$\frac{\partial S}{\partial U} = \frac{1}{T} = \frac{k}{\epsilon} \log \left(1 + \frac{\epsilon}{U}\right) \quad (26)$$

or

$$U = \epsilon/(e^{\epsilon/kT} - 1), \quad (27)$$

for the average energy of each radiator. The excellent agreement between (23) and experiment, together with Eq. (11), suggests that ϵ is inversely proportional to the wavelength or directly proportional to the frequency of the oscillator:

$$\epsilon = h\nu. \quad (28)$$

Then

$$U = h\nu/(e^{h\nu/kT} - 1), \quad (29)$$

and (11) implies

$$\rho(\nu) = (8\pi h\nu^3/c^3)/(e^{h\nu/kT} - 1) \quad (\text{Planck spectrum}), \quad (30)$$

for the spectral energy density of thermal radiation.

The expression (25) for S satisfies Eq. (22) with $A = k$ and $B = \epsilon$. Once (25) is obtained, therefore, one is led to the form (23) for the spectral energy density. The great success of (23) in fitting the experimental data led Planck to what he later called an "act of desperation" needed to derive (25).

One aspect of this desperate act is the way Planck counted the number of ways, or “complexions,” in which P energy elements could be distributed among N radiators. His counting procedure was totally at odds with classical statistical methods in its treatment of the energy elements as fundamentally *indistinguishable*. In one sense Planck was following Boltzmann in regarding all complexions as equally likely, but, of course, his way of counting the number of complexions was radically different and *might well be regarded as a precursor of Bose–Einstein statistics*.

Boltzmann had also employed energy elements in his counting of complexions, but in his calculations ϵ had no particular significance and, in fact, could ultimately be taken to be zero once a formula for W_N had been obtained. If Planck had taken the limit $\epsilon \rightarrow 0$ in Eq. (26), however, then $\partial S/\partial U \rightarrow k/U$ and $\partial^2 S/\partial U^2 \rightarrow -k/U^2$, which leads to the Rayleigh–Jeans distribution. In Planck’s derivation of his spectrum, therefore, the quantization of energy would appear to be essential.⁷

This is the traditional view of Planck’s innovation.⁸ It should be noted, however, that one of the main conclusions of Kuhn’s research² is that Planck did not in 1900 introduce any physical quantization of either radiation or material oscillators: “Planck’s concern ... had been and remained with *radiation*. His resonators were imaginary entities, not susceptible to experimental investigation. Their introduction was simply a device for bringing radiation to equilibrium, and it was justified, not by knowledge of the physical processes involved, but by Kirchhoff’s law, which made the equilibrium field independent of the equilibrium-producing material.”⁹

Until about 1905 Planck’s formula was regarded as little more than a superb fit to the experimental data. Its true significance began to be appreciated only when it was realized that the Rayleigh–Jeans law was an inevitable consequence of classical physics and the equipartition theorem, and therefore that the blackbody experiments had uncovered a fundamental failure of known (classical) theory.

A curious circumstance relating to zero-point energy, which was noted by Einstein and Stern,¹⁰ is worth mentioning. Consider the classical limit $kT \gg hv$ of the expression (29) for the average energy of an oscillator in thermal equilibrium with radiation:

$$U = \frac{hv}{e^{hv/kT} - 1} \cong \frac{hv}{1 + hv/kT + \frac{1}{2}(hv/kT)^2 - 1} = \frac{kT}{1 + \frac{1}{2}hv/kT} \cong kT - \frac{1}{2}hv. \quad (31)$$

Thus U does not reduce to kT , the energy predicted by the equipartition theorem, in the classical limit. But

$$U + \frac{1}{2}hv = \frac{hv}{e^{hv/kT} - 1} + \frac{1}{2}hv, \quad (32)$$

which includes the zero-point energy $\frac{1}{2}hv$, does reduce to kT in the classical limit. In Planck’s “second theory,” U was in fact replaced by $U + \frac{1}{2}hv$.

IV. PLANCK’S ZERO-POINT ENERGY

It was mentioned above that it took several years for the profound significance of Planck’s distribution to be appreciated. Planck himself was unsatisfied with the largely *ad hoc* theory he had used to derive his spectrum, and for

many years he explored alternative hypotheses that might lead to it.

In 1912, Planck published his “second theory.”¹¹ The absorption of radiation was assumed to proceed according to classical theory, whereas emission of radiation occurred discontinuously in discrete quanta of energy. Assume that an oscillator can radiate only after it has (continuously) absorbed an energy hv . Let P_n be the probability that it has energy between $(n-1)hv$ and nhv . When, as a result of absorption of radiation, its energy reaches nhv , there is a probability p that it will lose *all* its energy in the form of radiation and a probability $1-p$ that it continues to absorb without emission of radiation. Thus, $P_2 = P_1(1-p)$, $P_3 = P_2(1-p) = P_1(1-p)^2, \dots, P_n = P_1(1-p)^{n-1}$, and

$$\sum_{n=1}^{\infty} P_n = 1 = \sum_{n=1}^{\infty} P_1(1-p)^{n-1} = \frac{P_1}{p}, \quad (33)$$

or $P_1 = p$ is the probability that an oscillator in equilibrium with radiation has energy between 0 and hv , $P_2 = p(1-p)$ is the probability that it has energy between hv and $2hv$, and $P_n = p(1-p)^{n-1}$ is the probability that it has energy between $(n-1)hv$ and nhv . Following Boltzmann, Planck defines the oscillator entropy as

$$\begin{aligned} S &= -k \sum_{n=1}^{\infty} P_n \log P_n \\ &= -k \sum_{n=1}^{\infty} p(1-p)^{n-1} \log [p(1-p)^{n-1}] \\ &= -k \left[\frac{1}{p} \log p + \left(\frac{1}{p} - 1 \right) \log \left(\frac{1}{p} - 1 \right) \right]. \end{aligned} \quad (34)$$

Planck now assumes that all energies between $(n-1)hv$ and nhv are equally likely, so that the average energy of the oscillators with energy between $(n-1)hv$ and nhv is $\frac{1}{2}(n+n-1)hv = (n-\frac{1}{2})hv$. The average oscillator energy is then

$$\begin{aligned} U &= \sum_{n=1}^{\infty} \left(n - \frac{1}{2} \right) hv P_n = hv \sum_{n=1}^{\infty} \left(n - \frac{1}{2} \right) p(1-p)^{n-1} \\ &= (1/p - \frac{1}{2})hv \end{aligned} \quad (35)$$

or $1/p = U/hv + \frac{1}{2}$. From (34), therefore,

$$\begin{aligned} S &= k \left[\left(\frac{U}{hv} + \frac{1}{2} \right) \log \left(\frac{U}{hv} + \frac{1}{2} \right) \right. \\ &\quad \left. - \left(\frac{U}{hv} - \frac{1}{2} \right) \log \left(\frac{U}{hv} - \frac{1}{2} \right) \right]. \end{aligned} \quad (36)$$

Using once again the relation $\partial S/\partial U = 1/T$, Planck obtained

$$U = \frac{1}{2} hv \frac{e^{hv/kT} + 1}{e^{hv/kT} - 1} = \frac{hv}{e^{hv/kT} - 1} + \frac{1}{2} hv. \quad (37)$$

This implies that $U \neq 0$ when $T \rightarrow 0$: When $T \rightarrow 0$, $U \rightarrow \frac{1}{2}hv$. Planck’s equation (37) marked the birth of the concept of zero-point energy.

To derive $\rho(\nu)$, Planck could not resort to Eq. (11), since the derivation of that equation assumed continuous absorption and emission processes. Instead, he made the assumption that the ratio of the probability that an oscillator does not emit radiation, to the probability that it does, is proportional to $\rho(\nu):(1-p)/p = C\rho(\nu)$ or $1/p = C\rho(\nu) + 1$, where C is a constant of proportionality. This assumption is plausible in that, the greater the radiation

intensity, the more absorption should dominate emission. (Planck, of course, was not at this time aware of the possibility of *stimulated* emission!) Then, from (35), $U = [C\rho(\nu) + \frac{1}{2}]h\nu$ or

$$\rho(\nu) = \frac{1}{C} \frac{1}{e^{h\nu/kT} - 1}. \quad (38)$$

To determine C , Planck appeals to the classical limit, where the Rayleigh–Jeans law should apply: For $kT \gg h\nu$, $\rho(\nu)$ should reduce to (12), which requires that $1/C = 8\pi h\nu^3/c^3$ and therefore that

$$\rho(\nu) = \frac{8\pi h\nu^3/c^3}{e^{h\nu/kT} - 1}. \quad (39)$$

It is interesting that, in deducing C in this way, Planck was employing what would soon come to be called the “correspondence principle.” Furthermore, Planck’s probability p might well be regarded as the first example of a quantum transition probability.⁷

It is also noteworthy that in Planck’s second theory the material oscillators have zero-point energy, but the electromagnetic field does not: $\rho(\nu) \rightarrow 0$ for $T \rightarrow 0$. Had Planck simply used Eq. (11) to relate $\rho(\nu)$ and U , he would have obtained from (37) the spectral energy density

$$\rho'(\nu) = \rho(\nu) + \frac{4\pi h\nu^3}{c^3} = \frac{8\pi h\nu^3/c^3}{e^{h\nu/kT} - 1} + \frac{4\pi h\nu^3}{c^3}. \quad (40)$$

The zero-point energy appearing in Planck’s expression (37) is perfectly correct according to modern theory, even though Planck’s route to it is not. Furthermore, the zero-point electromagnetic energy appearing in (40) is also correct from the standpoint of modern quantum theory. By 1914, Planck was convinced that zero-point energy would be of no experimental consequence. However, the concept attracted much attention and soon came to play a major role in the work of Einstein.

V. THE EINSTEIN–HOPF MODEL

In his paper “Concerning a heuristic point of view toward the emission and transformation of light,” Einstein deduced that radiation satisfying the Wien distribution “behaves thermodynamically as though it consisted of a number of independent energy quanta of magnitude $[h\nu]$.”¹¹ Based on this viewpoint, he *predicted* the linear relation between radiation frequency and stopping potential in the photoelectric effect, a prediction confirmed by Millikan’s experiments in 1916. In 1906 he argued that “in emission and absorption the energy of a [Planck oscillator] changes by jumps which are integral multiples of $h\nu$.”¹² These were the beginnings of the photon concept.

Einstein struggled with the blackbody problem for more than 10 years after he introduced his heuristic viewpoint concerning energy quanta of radiation. In one important paper, Einstein and Hopf studied a simple model for the thermal equilibrium between oscillating dipoles and electromagnetic radiation.¹³ Imagine each dipole to consist of a particle of mass m and charge e , bound by an elastic restoring force to a mass M ($\gg m$) of opposite charge. The equation of motion for a linear dipole oscillator is then (see Appendix A)

$$\frac{d^2z}{dt^2} + \omega_0^2 z - \gamma \frac{d^3z}{dt^3} = \frac{e}{m} E_z(t), \quad (41)$$

where ω_0 ($= 2\pi\nu_0$) is the natural oscillation frequency,

$E_z(t)$ is the z component of the electric field acting on the particle, and $\gamma = 2e^2/3mc^3$. The two oppositely charged particles define an electric dipole moment $ez(t)$. Implicit in Eq. (41) is the *electric dipole approximation* of neglecting any spatial variation of $E_z(t)$ over the distance separating the particles. It is also assumed that the interaction of the dipole with the magnetic field is negligible.

Equation (41) is essentially the same equation used earlier by Planck to derive Eq. (11) (see Appendix A). In the Einstein–Hopf model, however, the dipole oscillators of mass $M + m \cong M$ are allowed to move; for simplicity they are constrained to move only along the x axis. Einstein and Hopf showed that there is a retarding force on a moving dipole as a result of its interaction with the field. This force acts to decrease its kinetic energy. Because of recoil associated with emission and absorption, however, the field also acts to increase the kinetic energy of a dipole. The condition for equilibrium is that the increase in kinetic energy due to recoil balances the decrease in kinetic energy associated with the retarding force.

Assuming $v/c \ll 1$, Einstein and Hopf showed that the retarding force due to motion through a thermal field of spectral energy density $\rho(\omega_0)$ is

$$F = -Rv, \quad (42)$$

where

$$R = \frac{4\pi^2 e^2}{5mc^2} \left[\rho(\omega_0) - \frac{\omega_0}{3} \frac{d\rho}{d\omega_0} \right], \quad (43)$$

and v is the velocity of the dipole. Essentially this same result is derived in Appendix B.

Consider now a dipole with linear momentum $Mv(t)$ at time t . After a short time τ , its momentum is

$$Mv(t + \tau) = Mv(t) + \Delta - Rv(t)\tau, \quad (44)$$

where Δ is the impulse imparted to the dipole in the time interval τ as a result of recoil associated with emission and absorption of radiation. Then,

$$M^2 v^2(t + \tau) - M^2 v^2(t) = \Delta^2 - 2MRv^2(t)\tau + (2M - R\tau)v(t)\Delta, \quad (45)$$

when M is taken to be large enough that terms quadratic in τ are negligible. Now, take the equilibrium ensemble average of both sides of (45):

$$2M \left[\left\langle \frac{1}{2} Mv^2(t + \tau) \right\rangle - \left\langle \frac{1}{2} Mv^2(t) \right\rangle \right] = 0 = \langle \Delta^2 \rangle - 4R\tau \left\langle \frac{1}{2} Mv^2(t) \right\rangle. \quad (46)$$

In writing this expression, we have used the fact that $\langle v(t)\Delta \rangle = 0$, since Δ is equally likely to be positive or negative in the time interval from t to $t + \tau$. In thermal equilibrium, furthermore, the equipartition theorem requires the average kinetic energy to be $\langle \frac{1}{2} Mv^2(t) \rangle = \frac{1}{2} kT$. The condition for thermal equilibrium is therefore

$$\tau^{-1} \langle \Delta^2 \rangle = 2RkT. \quad (47)$$

It remains to determine $\langle \Delta^2 \rangle$.

The force in the direction \hat{x} in which an electric dipole moment $\hat{z}p(t)$ is allowed to move in the Einstein–Hopf model is

$$F_x = p \frac{\partial E_x}{\partial z} + \frac{1}{c} (\hat{z}p \times \mathbf{B})_x = p \frac{\partial E_x}{\partial z} - \frac{1}{c} p B_y, \quad (48)$$

so that the impulse to the dipole during the time interval

from $t = 0$ to τ is

$$\begin{aligned} \Delta &= \int_0^\tau dt F_x = \int_0^\tau dt p \frac{\partial E_x}{\partial z} - \frac{1}{c} \int_0^\tau dt \dot{p} B_y \\ &= \int_0^\tau dt p \frac{\partial E_x}{\partial z} - \frac{1}{c} \int_0^\tau dt \left(\frac{\partial}{\partial t} (p B_y) - p \frac{\partial B_y}{\partial t} \right) \\ &= \int_0^\tau dt p \frac{\partial E_x}{\partial z} - \frac{1}{c} (p B_y) \Big|_0^\tau \\ &\quad - \int_0^\tau dt p \left(-\frac{\partial E_z}{\partial x} + \frac{\partial E_x}{\partial z} \right) \\ &= \int_0^\tau dt p \frac{\partial E_z}{\partial x} - \frac{1}{c} (p B_y) \Big|_0^\tau \\ &= e \int_0^\tau dt z(t) \frac{\partial E_z(t)}{\partial x} - \frac{e}{c} [z(t) B_y(t)] \Big|_0^\tau, \end{aligned} \quad (49)$$

where in the third line we have used the Maxwell equation $\nabla \times \mathbf{E} = -c^{-1} \partial \mathbf{B} / \partial t$. The second term in the last line of Eq. (49) is easily shown, using Eq. (51) below and a superposition similar to (50) for \mathbf{B} , to be independent of τ . Since τ is arbitrary, therefore, this term makes no contribution to $\langle \Delta^2 \rangle / \tau$ below and may be ignored for our purposes.

Einstein and Hopf write the electric field as a superposition of plane waves with independent random phases $\theta_{k\lambda}$:

$$\mathbf{E}(\mathbf{r}, t) = i \sum_{k\lambda} (A_{k\lambda} e^{-i(\omega_k t + \theta_{k\lambda})} - A_{k\lambda}^* e^{i(\omega_k t + \theta_{k\lambda})}) \mathbf{e}_{k\lambda}, \quad (50)$$

where $\mathbf{e}_{k\lambda}$ is a unit polarization vector for a plane wave with wave vector \mathbf{k} and linear polarization index λ ($= 1, 2$). A similar expression may be written for $\mathbf{B}(\mathbf{r}, t)$. The steady-state solution of Eq. (41) is then

$$z(t) = \frac{ie}{m} \sum_{k\lambda} (F_{k\lambda} e^{-i(\omega_k t + \theta_{k\lambda})} - F_{k\lambda}^* e^{i(\omega_k t + \theta_{k\lambda})}), \quad (51)$$

where the origin of coordinates has been chosen to be at the position of the dipole and $F_{k\lambda} = -A_{k\lambda} e_{k\lambda z} (\omega_k^2 - \omega_0^2 + i\gamma\omega_k^3)^{-1}$, where $e_{k\lambda z}$ is the z component of $\mathbf{e}_{k\lambda}$. In a separate paper, Einstein and Hopf show that $E_z(t)$ and $\partial E_z / \partial x$ must be treated as independent random variables in the time integral (49).¹⁴ It then follows from (49)–(51) by straightforward manipulations that $\langle \Delta \rangle = 0$ and

$$\tau^{-1} \langle \Delta^2 \rangle = (4\pi^4 c^4 \gamma / 5\omega_0^2) \rho^2(\omega_0), \quad (52)$$

where the ensemble average is taken over the random phases $\theta_{k\lambda}$.

Equation (47), together with (43) and (52), now gives a differential equation that must be satisfied by the spectral energy density of thermal radiation:

$$\rho(\omega) - \frac{\omega}{3} \frac{d\rho}{d\omega} = \left(\frac{\pi^2 c^3}{3\omega^2 kT} \right) \rho^2(\omega). \quad (53)$$

The solution of this equation satisfying $\rho(0) = 0$ is

$$\rho(\omega) = \omega^2 kT / \pi^2 c^3, \quad (54)$$

which is seen to be just the Rayleigh–Jeans law when we recall that $\omega = 2\pi\nu$ and $\rho(\omega) = \rho(\nu) / 2\pi$.

The beautifully cogent arguments of Einstein and Hopf provide further evidence that the Rayleigh–Jeans law is an inexorable consequence of classical physics. However, we shall see that their results are dramatically altered when zero-point energy is postulated.

VI. EINSTEIN AND STERN'S ZERO-POINT ENERGY

In 1913, Einstein and Stern noted that an *ad hoc* postulate about zero-point energy in the Einstein–Hopf model would lead to the Planck spectrum.¹⁰ First, let us note that Eq. (11) allows us to write (53) in a form in which the average dipole energy U appears explicitly:

$$\rho(\omega) - \frac{\omega}{3} \frac{d\rho}{d\omega} = \frac{1}{3kT} \rho(\omega) U. \quad (55)$$

Now, suppose the average oscillator energy U is replaced by $U + \hbar\omega$. This means that the dipole oscillators are now assumed to have a zero-point energy $\hbar\omega$. Equation (55) is then replaced by

$$\begin{aligned} \rho(\omega) - \frac{\omega}{3} \frac{d\rho}{d\omega} &= \frac{1}{3kT} \rho(\omega) U + \frac{\hbar\omega}{3kT} \rho(\omega) \\ &= \frac{\pi^2 c^3}{3\omega^2 kT} \rho^2(\omega) + \frac{\hbar\omega}{3kT} \rho(\omega) \\ &= \frac{\pi^2 c^3}{3\omega^2 kT} \left(\rho^2(\omega) + \frac{\hbar\omega^3}{\pi^2 c^3} \rho(\omega) \right). \end{aligned} \quad (56)$$

The solution of this equation satisfying $\rho(0) = 0$ is

$$\rho(\omega) = \frac{\hbar\omega^3 / \pi^2 c^3}{e^{\hbar\omega/kT} - 1} \quad (\text{Planck spectrum}). \quad (57)$$

In other words, if it is assumed that the dipole oscillators in the Einstein–Hopf model have a zero-point energy $\hbar\omega$, then the equilibrium spectrum of radiation is found to be the Planck spectrum.

The oscillator zero-point energy postulated by Einstein and Stern is twice that found earlier by Planck. Since we now know that Planck's zero-point energy $\frac{1}{2}\hbar\omega$ is the correct one, it is interesting to see how Einstein and Stern arrived at the correct spectrum using the wrong zero-point energy.

According to quantum theory, a field mode of frequency ω , like a material oscillator, has a zero-point energy $\frac{1}{2}\hbar\omega$. The total zero-point energy of a linear dipole oscillator of frequency ω and a field mode of the same frequency is therefore $\frac{1}{2}\hbar\omega + \frac{1}{2}\hbar\omega = \hbar\omega$. Einstein and Stern's zero-point energy $\hbar\omega$ is just this, *but they attributed it solely to the material dipole oscillators*.

Suppose we include in the Einstein–Hopf model a zero-point energy $\frac{1}{2}\hbar\omega$ for a dipole oscillator *and* a zero-point energy $\frac{1}{2}\hbar\omega$ for each field mode. Since there are $(8\pi\nu^2/c^3)d\nu = (\omega^2/\pi^2 c^3)d\omega$ field modes per unit volume in the frequency interval $[\omega, \omega + d\omega]$, the spectral energy density of the zero-point field is

$$\rho_0(\omega) = (\omega^2/\pi^2 c^3) \frac{1}{2}\hbar\omega = \hbar\omega^3 / 2\pi^2 c^3. \quad (58)$$

If we replace $\rho(\omega)$ in (55) by $\rho(\omega) + \rho_0(\omega)$, the left-hand side is unchanged:

$$\begin{aligned} [\rho(\omega) + \rho_0(\omega)] - (\omega/3) [\rho(\omega) + \rho_0(\omega)] \\ = \rho(\omega) - \frac{\omega}{3} \frac{d\rho}{d\omega}. \end{aligned} \quad (59)$$

If we also account for the zero-point energy of the dipole oscillators by replacing U by $U + \frac{1}{2}\hbar\omega$, the product $\rho(\omega)U$

on the right-hand side of (55) is changed to

$$\begin{aligned}
 & [\rho(\omega) + \rho_0(\omega)] (U + \frac{1}{2}\hbar\omega) \\
 &= \rho(\omega)U + \frac{1}{2}\hbar\omega\rho(\omega) + \rho_0(\omega)U + \frac{1}{2}\hbar\omega\rho_0(\omega) \\
 &= \frac{\pi^2c^3}{\omega^2} \left(\rho^2(\omega) + \rho_0(\omega)\rho(\omega) + \frac{\hbar\omega^3}{2\pi^2c^3}\rho(\omega) \right) \\
 &\quad + \frac{1}{2}\hbar\omega\rho_0(\omega) \\
 &= \frac{\pi^2c^3}{\omega^2} \left(\rho^2(\omega) + \frac{\hbar\omega^3}{\pi^2c^3}\rho(\omega) \right) + \frac{1}{2}\hbar\omega\rho_0(\omega), \quad (60)
 \end{aligned}$$

where we have used (58) and (11) in the form $U = (\pi^2c^3/\omega^2)\rho(\omega)$.

The term $\frac{1}{2}\hbar\omega\rho_0(\omega)$ in (60) results from a coupling of the zero-point motion of a dipole oscillator to the zero-point oscillations of the field. In quantum theory, in effect, no such coupling arises; an oscillator in its ground state in the absence of any applied field remains in its ground state. For our purposes here, let us just accept this result¹⁵ and drop the term $\frac{1}{2}\hbar\omega\rho_0(\omega)$ in (60):

$$\begin{aligned}
 & [\rho(\omega) + \rho_0(\omega)] (U + \frac{1}{2}\hbar\omega) \\
 &\rightarrow \frac{\pi^2c^3}{\omega^2} \left(\rho^2(\omega) + \frac{\hbar\omega^3}{\pi^2c^3}\rho(\omega) \right). \quad (61)
 \end{aligned}$$

From (54), (58), and (60), then, we have

$$\rho(\omega) - \frac{\omega}{3} \frac{d\rho}{d\omega} = \frac{\pi^2c^3}{3\omega^2kT} \left(\rho^2(\omega) + \frac{\hbar\omega^3}{\pi^2c^3}\rho(\omega) \right), \quad (62)$$

which is exactly the Einstein–Stern equation (56). The complete spectrum $\rho(\omega) + \rho_0(\omega)$ is then given by Eq. (40).

This route to the Planck spectrum may be summarized as follows. We modified the Einstein–Hopf model to include a zero-point energy $\frac{1}{2}\hbar\omega$ for a dipole oscillator and a zero-point energy $\frac{1}{2}\hbar\omega$ for each mode of the electromagnetic field, and anticipated a result of quantum theory that there is no contribution from the coupling of the zero-point oscillations of the dipole and the field. This led to the Einstein–Stern equation (56). Einstein and Stern, however, did not invoke any zero-point energy of the field, and to arrive at the Planck spectrum their dipole oscillators had to have a zero-point energy $\frac{1}{2}\hbar\omega$ plus what we know to be the zero-point energy of a field mode of the same frequency.

Why did Einstein and Stern not assume zero-point energy for the field? After all, one might have thought that the relation (11) between $\rho(\omega)$ and U would have made it obvious that, if either the dipole oscillator or the field has a zero-point energy, then so must the other. If Planck’s zero-point energy $\frac{1}{2}\hbar\omega$ is added to U in (11), for instance, then for consistency we must add the spectral energy density $\rho_0(\omega)$ of the zero-point field to $\rho(\omega)$:

$$\rho(\omega) + \rho_0(\omega) = (\omega^2/\pi^2c^3)(U + \frac{1}{2}\hbar\omega) \quad (63)$$

or again $\rho_0(\omega) = \hbar\omega^3/2\pi^2c^3$, which in turn implies that each field mode has a zero-point energy $\frac{1}{2}\hbar\omega$.

However, such a “consistency” argument rests on the usual acuity of hindsight. The fact is that at various stages in Einstein’s long efforts to understand the Planck spectrum he seriously doubted the general validity of Planck’s equation (11). This is not surprising, for if Planck had simply invoked equipartition of energy, and used $U = kT$ in (11), he would have obtained the Rayleigh–Jeans spectrum. It is not clear whether Planck was even aware at the time of the classical equipartition theorem. If he had known and believed the equipartition theorem, as Einstein later remarked, “he would probably not have made his great discovery.”¹⁶

There is another reason why Einstein and Stern might have been unwilling to attribute a zero-point energy to the field. If $\rho(\omega)$ and U are replaced by $\rho(\omega) + \rho_0(\omega)$ and $U + \frac{1}{2}\hbar\omega$, respectively, in the Einstein–Hopf model, then one obtains the Rayleigh–Jeans spectrum for the total spectral density $\rho(\omega) + \rho_0(\omega)$. Crucial to the derivation of the Planck spectrum is the omission of the term $\frac{1}{2}\hbar\omega\rho_0(\omega)$ in (59). This omission occurs automatically in the quantum theory of the Einstein–Hopf model.¹⁵ Without this consequence of quantum theory available to them, Einstein and Stern may have simply discounted the possibility of zero-point electromagnetic energy. Indeed, the first suggestion that there might be a zero-point electromagnetic field is due not to Planck or Einstein and Stern, but to Nernst.¹⁷

VII. EINSTEIN’S FLUCTUATION FORMULA

Prior to his work with Hopf and Stern, Einstein had derived a formula for the energy fluctuations of thermal radiation.¹⁸ Denoting the variance in energy in the volume V and in the frequency interval $[\omega, \omega + d\omega]$ by $\langle \Delta E_\omega^2 \rangle$, we may write the Einstein fluctuation formula as

$$\langle \Delta E_\omega^2 \rangle = \left(\hbar\omega\rho(\omega) + \frac{\pi^2c^3}{\omega^2}\rho^2(\omega) \right) V d\omega. \quad (64)$$

The importance of this formula lies in Einstein’s interpretation of it. The first term in brackets, according to Einstein, may be obtained “if radiation were to consist of independently moving pointlike quanta of energy $h\nu$ ”:

$$\langle \Delta E_\omega^2 \rangle_{\text{particles}} = \hbar\omega\rho(\omega) V d\omega, \quad (65)$$

whereas the second term follows when the field is treated as a superposition of independently fluctuating waves:

$$\langle \Delta E_\omega^2 \rangle_{\text{waves}} = (\pi^2c^3/\omega^2)\rho^2(\omega) V d\omega. \quad (66)$$

Thus, $\langle \Delta E_\omega^2 \rangle$ has both wave and particle contributions. The Einstein fluctuation formula was the earliest indicator of the wave–particle dualism in quantum theory.¹⁹

The “wave” term (66) may be derived from the superposition (50) of waves with independent random phases. For instance,

$$\begin{aligned}
 \langle \mathbf{E}^2(\mathbf{r}, t) \rangle &= -2 \operatorname{Re} \sum_{\mathbf{k}_1, \lambda_1} \sum_{\mathbf{k}_2, \lambda_2} \left\{ A_{\mathbf{k}_1, \lambda_1} A_{\mathbf{k}_2, \lambda_2} \exp [-i(\omega_{\mathbf{k}_1} + \omega_{\mathbf{k}_2})t] e^{i(\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r}} \langle \exp [-i(\theta_{\mathbf{k}_1, \lambda_1} + \theta_{\mathbf{k}_2, \lambda_2})] \rangle \right. \\
 &\quad \left. - A_{\mathbf{k}_1, \lambda_1} A_{\mathbf{k}_2, \lambda_2}^* \exp [-i(\omega_{\mathbf{k}_1} - \omega_{\mathbf{k}_2})t] e^{i(\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{r}} \langle \exp [-i(\theta_{\mathbf{k}_1, \lambda_1} - \theta_{\mathbf{k}_2, \lambda_2})] \rangle \right\} \mathbf{e}_{\mathbf{k}_1, \lambda_1} \cdot \mathbf{e}_{\mathbf{k}_2, \lambda_2}, \quad (67)
 \end{aligned}$$

where again the average is over the phases $\theta_{\mathbf{k}\lambda}$, which are assumed to be independent, uniformly distributed random variables on the interval $[0, 2\pi]$. Thus

$$\langle \mathbf{E}^2(\mathbf{r}, t) \rangle = 2 \sum_{\mathbf{k}\lambda} |A_{\mathbf{k}\lambda}|^2 \quad (68)$$

and, similarly,

$$\langle E^4(\mathbf{r},t) \rangle = 8 \left(\sum_{\mathbf{k}\lambda} |A_{\mathbf{k}\lambda}|^2 \right)^2, \quad (69)$$

so that

$$\langle E^4(\mathbf{r},t) \rangle - \langle E^2(\mathbf{r},t) \rangle^2 = 4 \left(\sum_{\mathbf{k}\lambda} |A_{\mathbf{k}\lambda}|^2 \right)^2 = \langle E^2(\mathbf{r},t) \rangle^2. \quad (70)$$

Since the electromagnetic energy density is proportional to $\langle E^2 \rangle$, it follows from (70) that the variance in energy associated with frequency ω is proportional to $\rho^2(\omega)$. We omit the trivial details of the derivation, which leads directly to Eq. (66).

The “particle” term (65) in the Einstein fluctuation formula is of far less obvious origin, and to derive it we temporarily assume the field energy can be written as

$$E = \sum_{\mathbf{k}\lambda} n_{\mathbf{k}\lambda} \hbar \omega_{\mathbf{k}}, \quad (71)$$

so that its variance is

$$\langle \Delta E^2 \rangle = \sum_{\mathbf{k}\lambda} \langle \Delta n_{\mathbf{k}\lambda} \rangle^2 \hbar^2 \omega_{\mathbf{k}}^2, \quad (72)$$

where the $n_{\mathbf{k}\lambda}$ are integers. Thus, we are assuming that the field energy is composed of discrete quanta of energy $\hbar \omega_{\mathbf{k}}$ and that the numbers of quanta associated with different modes fluctuate independently. We assume Poisson statistics for these quanta, so that

$$\langle \Delta n_{\mathbf{k}\lambda}^2 \rangle = \langle n_{\mathbf{k}\lambda} \rangle \quad (73)$$

and

$$\langle \Delta E^2 \rangle = \sum_{\mathbf{k}\lambda} \langle n_{\mathbf{k}\lambda} \rangle \hbar^2 \omega_{\mathbf{k}}^2. \quad (74)$$

Since $\rho(\omega)$ is proportional to the average number of photons at frequency ω , Eq. (74) leads easily to the particle term (65) in the Einstein fluctuation formula.

The Einstein fluctuation formula is easily derived more thoroughly.¹⁹ For the present discussion, we simply note that we can obtain both the “wave” and “particle” terms using the classical wave picture with zero-point energy. That is, if we replace $\rho(\omega)$ in (66) by $\rho(\omega) + \rho_0(\omega)$, where the spectral energy density $\rho_0(\omega)$ of the zero-point field is given by (57), we have

$$\begin{aligned} \langle \Delta E_\omega^2 \rangle_{\text{waves}} &\rightarrow \frac{\pi^2 c^3}{\omega^2} [\rho^2(\omega) + 2\rho(\omega_0)\rho(\omega) + \rho_0^2(\omega)] V d\omega \\ &= \left(\frac{\pi^2 c^3}{\omega^2} \rho^2(\omega) + \hbar \omega \rho(\omega) \right) V d\omega + \frac{\pi^2 c^3}{\omega^2} \rho_0^2(\omega) V d\omega \\ &= \langle \Delta E_\omega^2 \rangle_{\text{waves}} + \langle \Delta E_\omega^2 \rangle_{\text{particles}} + \frac{\pi^2 c^3}{\omega^2} \rho_0^2(\omega) V d\omega \\ &= \langle \Delta E_\omega^2 \rangle_{\text{waves}} + \langle \Delta E_\omega^2 \rangle_{\text{particles}} + \frac{1}{2} \hbar \omega \rho_0(\omega) V d\omega. \end{aligned} \quad (75)$$

The “extra” (third) term in this expression does not appear in the Einstein fluctuation formula. Indeed, it does not appear at all in quantum theory, for the same reason that the term $\frac{1}{2} \hbar \omega \rho_0(\omega)$ in Eq. (60) is absent in quantum theory.

But aside from this spurious “extra” term, we have obtained the Einstein fluctuation formula from a classical wave perspective that includes zero-point field energy. Obviously, the argument is essentially the same as in our approach to the Einstein–Stern theory and suggests that *the particle term in the Einstein fluctuation formula may be regarded as a consequence of zero-point field energy.*

The particle term was in fact the novel element in Einstein’s fluctuation formula, and Einstein emphasized that this term was incompatible with classical wave theory. If there were only classical wave fluctuations in thermal radiation, we could ignore the term proportional to $\rho(\omega)$ in Eq. (62). The result is

$$\rho(\omega) - \frac{\omega}{3} \frac{d\rho}{d\omega} = \frac{\pi^2 c^3}{3\omega^2 kT} \rho^2(\omega), \quad (76)$$

and the solution is the Rayleigh–Jeans spectrum $\rho(\omega) = (\omega^2/\pi^2 c^3) kT$. Without the wave term, on the other hand, (62) becomes

$$\rho(\omega) - \frac{\omega}{3} \frac{d\rho}{d\omega} = \frac{\hbar \omega}{3kT} \rho(\omega), \quad (77)$$

and the solution of this equation is the Wien distribution $\rho(\omega) = (\hbar \omega^3/\pi^2 c^3) e^{-\hbar \omega/kT}$. This is consistent with the

fact that in 1905 Einstein had deduced his “heuristic point of view” concerning radiation energy quanta by considering only radiation satisfying the Wien distribution.¹¹

VIII. EINSTEIN’S A AND B COEFFICIENTS

Einstein wrote to his friend Besso in November 1916 that “A splendid light has dawned on me about the absorption and emission of radiation.” He was referring to his new insight into the “heuristic principle” of 1905 and the basis it provided for an “astonishingly simple” derivation of the Planck spectrum.²⁰

For the sake of completeness, we summarize the argument here. Einstein assumes that an atom (or molecule) has discrete energy levels. Let N_1 and N_2 be the numbers of atoms in energy levels E_1 and E_2 , respectively, with $E_2 > E_1$. (For simplicity we ignore the possibility of level degeneracies, which does not affect the result for the spectral density of thermal radiation.) The rate at which N_1 changes due to the absorption of radiation, with the atom making an upward transition to the level E_2 , is assumed to be proportional to N_1 and the spectral energy density $\rho(\omega_0)$ at the Bohr transition frequency $\omega_0 = (E_2 - E_1)/\hbar$:

$$(\dot{N}_1)_{\text{absorption}} = -B_{12} N_1 \rho(\omega_0). \quad (78)$$

Einstein proposes two kinds of emission processes by which an atom can jump from level E_2 to E_1 with the emission of radiation of frequency ω_0 . One is *spontaneous* emis-

sion, which can occur in the absence of any radiation and is described by the rate constant A_{21} :

$$(\dot{N}_1)_{\text{spontaneous emission}} = A_{21}N_2. \quad (79)$$

The other is *stimulated* emission, which is assumed to proceed at a rate proportional to both N_2 and $\rho(\omega_0)$:

$$(N_1)_{\text{stimulated emission}} = B_{21}N_2\rho(\omega_0). \quad (80)$$

The condition for equilibrium is

$$(\dot{N}_1)_{\text{absorption}} + (\dot{N}_1)_{\text{spontaneous emission}} + (\dot{N}_1)_{\text{stimulated emission}} = 0 \quad (81)$$

or

$$A_{21}N_2 + B_{21}N_2\rho(\omega_0) = B_{12}N_1\rho(\omega_0), \quad (82)$$

$$\begin{aligned} \rho(\omega_0) &= \frac{A_{21}/B_{21}}{(B_{12}/B_{21})(N_1/N_2) - 1} \\ &= \frac{A_{21}/B_{21}}{(B_{12}/B_{21})e^{\hbar\omega_0/kT} - 1}, \end{aligned} \quad (83)$$

since $N_2/N_1 = e^{-(E_2 - E_1)/kT} = e^{-\hbar\omega_0/kT}$ in thermal equilibrium. We are using Bohr's postulate that $E_2 - E_1 = \hbar\omega_0$, but it is worth noting that this relation in fact emerged naturally from Einstein's analysis once the assumption of discrete energy levels was made and the Wien displacement law was invoked.

At very high temperatures, $\rho(\omega_0)$ becomes so large that spontaneous emission is much less probable than stimulated emission. Then, from (82) we must have $B_{21} = B_{12}$ and, from (83),

$$\rho(\omega_0) = (A_{21}/B_{21})/(e^{\hbar\omega_0/kT} - 1). \quad (84)$$

For $kT \gg \hbar\omega_0$, furthermore,

$$\rho(\omega_0) \cong (A_{21}/B_{21})(kT/\hbar\omega_0). \quad (85)$$

This is the limit where the radiation energy quanta become so small compared with kT that the classical Rayleigh-Jeans law should be applicable. This requires $(A_{21}/B_{21})(kT/\hbar\omega_0) = (\omega_0^2/\pi^2c^3)kT$ or

$$A_{21}/B_{21} = \hbar\omega_0^3/\pi^2c^3, \quad (86)$$

and Eq. (84) then yields the Planck spectrum for $\rho(\omega)$.

This derivation of the Planck spectrum combined aspects of Einstein's earlier work on radiation quanta with the theories of Planck and Bohr. But in it Einstein had made several profoundly important theoretical advances, and he suggested that "The simplicity of the hypotheses makes it seem probable...that these will become the basis of the future theoretical description." He was absolutely correct. None of the developments since 1917 have required any modification of Einstein's derivation of the blackbody spectrum.

One major consequence of Einstein's work, of course, was the introduction of the concept of stimulated emission. Without the stimulated emission term, (82) and (83) are replaced by

$$A_{21}N_2 = B_{12}N_1\rho(\omega_0), \quad (87)$$

$$\rho(\omega_0) = \frac{A_{21}}{B_{12}} \frac{N_2}{N_1} = \frac{\hbar\omega_0^3}{\pi^2c^3} e^{-\hbar\omega_0/kT}. \quad (88)$$

Without stimulated emission, therefore, Einstein would have obtained the Wien distribution.

Einstein's work was also the first to reveal atomic radiation in the form of spontaneous emission as a nonclassical

process in which "God plays dice": There is nothing to tell us exactly *when* the atom will make a spontaneous jump to a state of lower energy. Einstein later wrote to Born that "That business about causality causes me a lot of trouble... Can the quantum absorption and emission of light ever be understood in the sense of the complete causality requirement, or would a statistical residue remain?... I would be very unhappy to renounce complete causality."⁷ That displeasure prevented Einstein from ever accepting quantum theory as a complete description of Nature.

Another novel aspect of Einstein's work was that it brought out the fact that photons carry linear momentum $h\nu/c$ as well as energy $h\nu$. This part of Einstein's work of 1917 is not nearly as widely known as the derivation of the Planck spectrum just reviewed. According to Einstein, however, "a theory [of thermal radiation] can only be regarded as justified when it is able to show that the impulses transmitted by the radiation field to matter lead to motions that are in accordance with the theory of heat." Einstein showed that the momentum transfers accompanying emission and absorption are consistent with statistical mechanics *if the thermal radiation follows the Planck distribution*.

Consider the interaction with radiation of an atom initially at rest in the laboratory frame of reference. After a time τ , it acquires some linear momentum Δ as a result of emission and absorption of radiation. Each emission or absorption process imparts to the atom a linear momentum λ_i , which may be positive or negative. If n emission and absorption processes occur during the time interval τ , then

$$\Delta = \sum_{i=1}^n \lambda_i, \quad (89)$$

and, assuming the λ_i to be independent random variables of zero mean,

$$\langle \Delta^2 \rangle = \sum_{i=1}^n \langle \lambda_i^2 \rangle \rightarrow \frac{1}{3} \left(\frac{\hbar\omega_0}{c} \right)^2 n \quad (90)$$

if we associate with each process of emission or absorption a momentum transfer (photon momentum) $\hbar\omega_0/c$. We have also included a factor of $\frac{1}{3}$ because, as in the Einstein-Hopf model, the atoms are assumed to move in only one direction. The average number n of emission and absorption events occurring in the time interval τ is given, according to the foregoing analysis, by

$$n = N_2 A_{21} \tau + (N_1 + N_2) B_{12} \rho(\omega_0) \tau, \quad (91)$$

so that

$$\begin{aligned} \tau^{-1} \langle \Delta^2 \rangle &= \frac{1}{3} \left(\frac{\hbar\omega_0}{c} \right)^2 [N_2 A_{21} + (N_1 + N_2) B_{12} \rho(\omega_0)] \\ &= \frac{2}{3} \left(\frac{\hbar\omega_0}{c} \right)^2 N_1 B_{12} \rho(\omega_0), \end{aligned} \quad (92)$$

where we have used the equilibrium condition (82).

This result shows that an atom interacting with radiation will continually gain kinetic energy unless there is some retarding force to maintain the fixed average kinetic energy $\langle \frac{1}{2}mv^2 \rangle = \frac{1}{2}kT$ demanded by statistical mechanics. The origin of this retarding force is the same as in the Einstein-Hopf model, except that now we must express it in terms of quantities characteristic of an *atom* rather than a classical dipole oscillator. As shown in Appendix B, this

force is given by the formula

$$F = -Rv = -\left(\frac{\hbar\omega_0}{c^2}\right)(N_1 - N_2) \times B_{12}\left(\rho(\omega_0) - \frac{\omega_0}{3} \frac{d\rho}{d\omega_0}\right)v. \quad (93)$$

As in the classical Einstein-Hopf model, the condition for thermal equilibrium is $\langle \Delta^2 \rangle / \tau = 2RkT$ or, from (92) and (93),

$$\rho(\omega_0) - \frac{\omega_0}{3} \frac{d\rho}{d\omega_0} = \left(\frac{\hbar\omega_0}{3kT}\right)\left(\frac{N_1}{N_1 - N_2}\right)\rho(\omega_0) = \left(\frac{\hbar\omega_0/3kT}{1 - e^{-\hbar\omega_0/kT}}\right)\rho(\omega_0). \quad (94)$$

The solution of this equation is the Planck spectrum. Thus, Einstein showed that in his theory of thermal radiation, "the impulses transmitted by the radiation field to matter lead to motions that are in accordance with the theory of heat."

IX. ZERO-POINT ENERGY, SPONTANEOUS EMISSION, AND BOSE-EINSTEIN STATISTICS

In Sec. VI we alluded to the fact that an oscillator (or atom) in its ground state does not absorb zero-point electromagnetic radiation. The question arises whether an excited atom undergoes *stimulated emission* due to the zero-point field.

Let us suppose that it does. Then, according to the Einstein theory described in Sec. VIII, the rate at which an atom in level 2 is stimulated by the zero-point field to drop to level 1 should be given by

$$(\dot{N}_2)_{\text{stimulated emission}}^0 = -B_{21}\rho_0(\omega_0)N_2 = -B_{21}(\hbar\omega_0^3/2\pi^2c^3)N_2, \quad (95)$$

where we have used Eq. (58) for the spectral energy density $\rho_0(\omega)$ of the zero-point field. Using (86), therefore, we infer

$$(\dot{N}_2)_{\text{stimulated emission}}^0 = -B_{21}\left(\frac{A_{21}}{2B_{21}}\right)N_2 = -\frac{1}{2}A_{21}N_2 = \frac{1}{2}(\dot{N}_2)_{\text{spontaneous emission}}. \quad (96)$$

Thus, we can almost interpret spontaneous emission as stimulated emission due to the zero-point field—*almost* because we calculate within this interpretation only half the correct A coefficient for spontaneous emission. In spite of this discrepancy, one repeatedly hears and reads statements to the effect that "spontaneous emission appears as a forced emission caused by the zero-point oscillations of the electromagnetic field."²¹

The result (96), however, does suggest that spontaneous emission has something to do with zero-point radiation, even if it is not simply emission induced by this radiation. Another way to infer this is to use the equation

$$\frac{N_1}{N_1 - N_2} = 1 + \frac{B_{21}}{A_{21}}\rho(\omega_0), \quad (97)$$

which follows from (81), in Eq. (94):

$$\rho(\omega_0) - \frac{\omega_0}{3} \frac{d\rho}{d\omega_0} = \frac{\hbar\omega_0}{3kT}\left(1 + \frac{B_{21}}{A_{21}}\rho(\omega_0)\right)\rho(\omega_0) = \frac{\pi^2c^3}{3\omega_0^2kT}\left(\rho^2(\omega_0) + \frac{A_{21}}{B_{21}}\rho(\omega_0)\right). \quad (98)$$

The identity (86) shows that this result is equivalent to (62). But now it is evident that the second term in brackets is associated with spontaneous emission. In other words, *the particle term in the Einstein fluctuation formula is related to spontaneous emission*. The fact that the particle term may also be related as in Sec. VII to the zero-point field thus suggests again some connection between spontaneous emission and the zero-point field.²²

Earlier, we remarked that Planck's "desperate" method of counting energy elements was a foreshadowing of Bose-Einstein statistics. A connection between zero-point energy and Bose-Einstein statistics is revealed by Einstein's fluctuation formula. Recall first that Bose-Einstein statistics imply the variance

$$\langle [\Delta n(\omega)]^2 \rangle = \langle n(\omega) \rangle + \langle n(\omega) \rangle^2, \quad (99)$$

at thermal equilibrium, where $\langle n(\omega) \rangle$ is the average number of photons at frequency ω . Since $\rho(\omega) = (\hbar\omega^3/\pi^2c^3)\langle n(\omega) \rangle$, (99) implies the Einstein-fluctuation formula (64). Based on our discussion in Sec. VII, we can infer that the term $\langle n(\omega) \rangle$ in (99) is attributable to zero-point energy and is closely connected, furthermore, with the possibility of spontaneous emission. This has been discussed in more detail elsewhere.¹⁵

X. SPECIFIC HEATS

The specific heat of a solid will in general have contributions from both electronic and vibrational degrees of freedom. Except at very high temperatures, however, the electrons are all in their ground states and make no contribution to the specific heat. Then, the N atoms making up the solid may be regarded as inert vibrators, and under the approximation of harmonic vibrations the total energy for the $3N$ degrees of freedom is $U = 3NkT$. Thus, $dU/dT = 3Nk$, and the specific heat per mole is

$$c_v = 3N_A k = 3R \approx \text{cal/mol K} \quad (\text{Dulong-Petit law}). \quad (100)$$

This classical prediction is the Dulong-Petit law, named after the experimenters who observed it in 1819 for 12 metals and sulfur at room temperature. As the temperature is decreased, however, c_v is found to decrease, and $c_v \rightarrow 0$ as $T \rightarrow 0$, contradicting the classical prediction (100) based on the equipartition theorem.

It was found in 1840 that the specific heat of diamond is smaller than 6 cal/(mol K) even at room temperature. This anomaly was first explained by Einstein in 1907.²³ Einstein argued that Planck's equation (29) gives the average energy in thermal equilibrium of each (harmonic) vibrational degree of freedom, so that

$$U = 3Nh\nu/(e^{h\nu/kT} - 1) \quad (101)$$

and

$$c_v = 3R(\theta/T)^2[e^{\theta/T}/(e^{\theta/T} - 1)^2] \quad (102)$$

is the specific heat per mole, where $\theta \equiv h\nu/k$ is the "Einstein temperature," the one adjustable parameter in Einstein's theory. For high temperatures ($T \gg \theta$), Eq. (102)

reduces to the Dulong–Petit law. At low temperatures, however, c_v is less than the Dulong–Petit value, and in particular $c_v \rightarrow 0$ as $T \rightarrow 0$. From a fit to experimental data, Einstein deduced that $\theta \approx 1300$ K for diamond. A substance with such a large value of θ will have a small value of c_v even at room temperature.

In 1913, Einstein and Stern, in the same paper discussed in connection with the blackbody problem in Sec. VI, turned their attention to the specific heats of gases. Their work was motivated by the recent report by Eucken²⁴ that the molar specific heat for H_2 at room temperature was about 5 cal/(mol K) at room temperature, but about 3 cal/(mol K) at $T \approx 60$ K. Einstein and Stern suggested that this behavior was a consequence of molecular rotations and zero-point energy.

The energy of a dumbbell rotator with moment of inertia I and rotational frequency ν is $\frac{1}{2}I(2\pi\nu)^2$. Suppose, following Einstein and Stern, that in thermal equilibrium this energy is given by the Planck equation (29):

$$U = \frac{1}{2}I(2\pi\nu)^2 = h\nu/(e^{h\nu/kT} - 1). \quad (103)$$

The rotational contribution to the specific heat is then

$$\begin{aligned} c_r &= N_A \frac{dU}{dT} = N_A \frac{dU}{d\nu} \frac{d\nu}{dT} \\ &= N_A (2\pi^2 I \nu) \frac{d\nu}{dT} \\ &= \left(\frac{R}{k}\right) p \nu \frac{d\nu}{dT}, \end{aligned} \quad (104)$$

where $p \equiv 2\pi^2 I$. From Eq. (103) it is clear that ν is a function of T ; $d\nu/dT$ follows by differentiation of both sides of that equation with respect to T ,

$$\frac{d\nu}{dT} = \frac{\nu}{T} \left(1 + \frac{kT}{p\nu^2 + h\nu}\right)^{-1}, \quad (105)$$

and it follows from (104) that

$$c_r = R \frac{2p\nu^2}{kT} \left(1 + \frac{kT}{p\nu^2 + h\nu}\right)^{-1}, \quad (106)$$

where $\nu(T)$ is found by solution of (103). The rotational specific heat calculated in this way for the example $p = 2.9 \times 10^{-40}$ g cm² considered by Einstein and Stern is shown in Fig. 1. The predicted dependence of the specific heat on temperature is quite different from the dependence observed by Eucken, and in particular the predicted specific heats at low temperatures are much too large.

Now suppose, however, that Eq. (103) is modified to include zero-point energy:

$$U = p\nu^2 = h\nu/(e^{h\nu/kT} - 1) + \frac{1}{2}h\nu. \quad (107)$$

Following the same steps leading from (103) to (106), it is found that

$$c_r = R \frac{2p\nu^2}{kT} \left(1 + \frac{kT}{p\nu^2 - h^2/4p}\right)^{-1}, \quad (108)$$

where $\nu(T)$ is obtained by solving (107) for ν in terms of T . The resulting c_r plotted in Fig. 1 is seen to agree very well with Eucken's observations. At high temperatures c_r asymptotes to $R \approx 2$ cal/(mol K), but at low temperatures $c_r \rightarrow 0$ at much larger temperatures than predicted by Eq. (106).

Einstein and Stern thus gave a very interesting interpre-

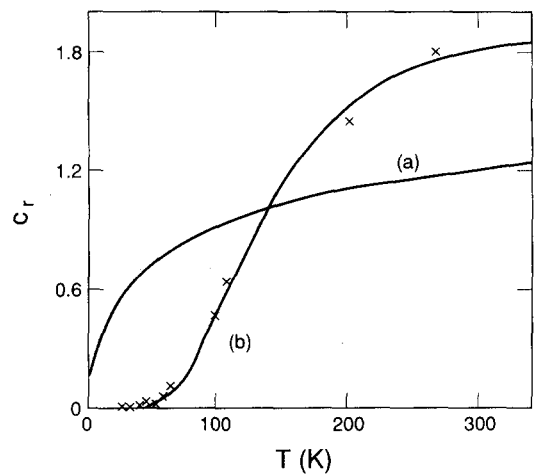


Fig. 1. Experimental data (×) of Eucken (Ref. 24) on specific heat of molecular hydrogen; specific heat computed by Einstein and Stern (a) without zero-point energy [Eq. (106)] and (b) with zero-point energy [Eq. (108)].

tation of Eucken's observation that the specific heat of H_2 decreased from 5 to 3 cal/(mol K) as T decreased from 300 to 60 K: Because of zero-point energy, the rotational contribution to the specific heat decreases from 2 to 0 cal/(mol K) as T decreases. That is, the existence of zero-point energy causes the rotational specific heat of a gas to "freeze out." Einstein and Stern concluded that "The existence of a zero-point energy of size $\frac{1}{2}h\nu$ [is] is probable."

The Einstein–Stern explanation turned out to be incorrect. The quantum-mechanical rotational energy levels of a diatomic molecule are given by $E_J \approx BJ(J+1)$, where B is a constant characteristic of the molecule and $J = 0, 1, 2, \dots$. There, a molecule has no zero-point rotational energy. On the other hand, Einstein and Stern were correct in their hypothesis that the observed decrease of specific heat with temperature of H_2 was connected to molecular rotations.

According to quantum mechanics, the fact that c_v and $c_r \rightarrow 0$ as $T \rightarrow 0$ is due simply to the fact that there are discrete energy levels associated with the internal degrees of freedom of a molecule. If kT is small compared with the energy separation between the lowest and first-excited energy levels, there is a high probability that only the lowest-energy state is occupied, and so the specific heat corresponding to that degree of freedom is "frozen out" in the sense that dU/dT decreases with T and approaches zero as $T \rightarrow 0$.

XI. X-RAY DIFFRACTION

An important question, prior to the first experiments, was whether x-ray diffraction would be spoiled by the thermal motions of the atoms in crystal lattices. It was first shown by Debye in 1914 that these thermal motions basically just reduce the intensity of a diffracted beam from that predicted for an idealized lattice of stationary atoms.²⁵ Debye also showed that if Planck's zero-point energy were real, then there should be such a reduction in intensity even as $T \rightarrow 0$. We now know that zero-point motion can indeed have a significant effect on x-ray diffraction. In this section we will briefly sketch a derivation of the so-called Debye–Waller factor that accounts for the motion of lattice atoms.

Consider the field far from a collection of identical scatterers. We assume the n th scatterer at \mathbf{r}_n has strength p_n ,

and write the total scattered field at \mathbf{r} as

$$E_s(\mathbf{r}) = \sum_n \frac{p_n}{|\mathbf{r}-\mathbf{r}_n|} e^{-i\omega(t-|\mathbf{r}-\mathbf{r}_n|)/c}$$

$$= e^{-i\omega t} \sum_n \frac{p_n}{|\mathbf{r}-\mathbf{r}_n|} e^{ik|\mathbf{r}-\mathbf{r}_n|} \quad (109)$$

For distances large compared with the dimensions of the scattering volume,

$$|\mathbf{r}-\mathbf{r}_n| = (r^2 - 2\mathbf{r}\cdot\mathbf{r}_n + r_n^2)^{1/2}$$

$$= r(1 - 2\mathbf{r}\cdot\mathbf{r}_n/r^2 + r_n^2/r^2)^{1/2}$$

$$\cong r(1 - \mathbf{r}\cdot\mathbf{r}_n/r^2) = r - \mathbf{r}\cdot\mathbf{r}_n/r, \quad (110)$$

so that

$$k|\mathbf{r}-\mathbf{r}_n| \cong kr - (k\mathbf{r}/r)\cdot\mathbf{r}_n \cong kr - \mathbf{k}\cdot\mathbf{r}_n,$$

in the exponential in (109), where \mathbf{k} is the wave vector of the (elastically) scattered wave. Thus

$$E_s(\mathbf{r}) \cong \frac{1}{r} e^{-i\omega(t-r/c)} \sum_n p_n e^{-i\mathbf{k}\cdot\mathbf{r}_n} \quad (111)$$

We take the strength p_n of the n th scatterer to be proportional to the field $E_0 e^{i\mathbf{k}_0\cdot\mathbf{r}_n}$ ($k_0 = k$) incident upon it: $p_n = \alpha E_0 e^{i\mathbf{k}_0\cdot\mathbf{r}_n}$ and

$$E_s(\mathbf{r}) \cong \frac{\alpha}{r} e^{-i(\omega t - kr)} \sum_n e^{-i\mathbf{K}\cdot\mathbf{r}_n}, \quad (112)$$

where $\mathbf{K} \equiv \mathbf{k} - \mathbf{k}_0$.

For a periodic lattice of scatterers, the scattered field (112) is nonvanishing only in directions such that \mathbf{K} belongs to the reciprocal lattice. For a one-dimensional lattice, for instance, this means that $Kd = 2\pi n$, where d is the lattice spacing and n is an integer. Since

$$K = (k^2 + k_0^2 - 2\mathbf{k}_0\cdot\mathbf{k})^{1/2}$$

$$= (2k^2 - 2k^2 \cos 2\theta)^{1/2}$$

$$= 2k \sin \theta = (4\pi/\lambda) \sin \theta,$$

the condition that K belongs to the reciprocal lattice is just the Bragg condition, $2d \sin \theta = n\lambda$, where 2θ is the angle between the incident and scattered (diffracted) waves.

Now, let us take into account the thermal motion of the atoms, replacing \mathbf{r}_n above by $\mathbf{r}_n + \mathbf{u}$, where \mathbf{u} represents a displacement from a fixed lattice site. Then,

$$\sum_n e^{-i\mathbf{K}\cdot\mathbf{r}_n} \rightarrow e^{-i\mathbf{K}\cdot\mathbf{u}} \sum_n e^{-i\mathbf{K}\cdot\mathbf{r}_n} \quad (113)$$

We are interested in the average of $e^{-i\mathbf{K}\cdot\mathbf{u}}$ as \mathbf{u} undergoes thermal motion:

$$\langle e^{-i\mathbf{K}\cdot\mathbf{u}} \rangle = 1 - i\mathbf{K}\cdot\langle\mathbf{u}\rangle - \frac{1}{2}\langle(\mathbf{K}\cdot\mathbf{u})^2\rangle + \dots$$

$$= 1 - \frac{1}{6}K^2\langle\mathbf{u}^2\rangle + \dots, \quad (114)$$

since $\langle\mathbf{u}\rangle = 0$. The two terms shown explicitly are the first two terms of the Taylor series for $\exp[-(1/6)K^2\langle\mathbf{u}^2\rangle]$. In fact, if the oscillations of \mathbf{u} are assumed to be harmonic, we have

$$\langle e^{-i\mathbf{K}\cdot\mathbf{u}} \rangle = e^{-(1/6)K^2\langle\mathbf{u}^2\rangle}, \quad (115)$$

and $\frac{1}{2}m\omega_0^2\langle\mathbf{u}^2\rangle = \frac{3}{2}kT$, where m and ω_0 are the mass and frequency of the harmonic oscillations; for simplicity we assume the elastic restoring force is the same in all directions. Thus the thermal fluctuations in the atomic positions cause the diffracted beam to be reduced in intensity by the

factor

$$|\langle e^{-i\mathbf{K}\cdot\mathbf{u}} \rangle|^2 \equiv e^{-2W} = e^{-K^2 kT/2m\omega_0^2}. \quad (116)$$

This is called the *Debye-Waller factor*. Our classical hand-waving derivation gives the correct order of magnitude for this factor.

But the classical model of lattice vibrations breaks down, of course, at low temperatures. In particular, as $T \rightarrow 0$ there is a nonvanishing $\langle\mathbf{u}^2\rangle$ associated with zero-point energy:

$$\frac{1}{2}m\omega_0^2\langle\mathbf{u}^2\rangle = 3\left(\frac{1}{2}\hbar\omega_0\right), \quad (117)$$

so that

$$e^{-2W} = e^{-\hbar K^2/m\omega_0}, \quad \text{for } T \rightarrow 0. \quad (118)$$

This gives the correct order of magnitude for the zero-temperature Debye-Waller factor.

XII. MOLECULAR VIBRATIONS

Direct evidence for the reality of zero-point energy was provided by Mulliken in 1924.²⁶ Consider the vibrational spectra of two diatomic molecules differing only by having different nuclear isotopes. The masses of these two vibrators are then different and, consequently, so are their vibrational frequencies. For relatively heavy molecules, these differences are small but readily observable. According to quantum mechanics, each molecule has vibrational energy levels given by $E_n = (n + \frac{1}{2})\hbar\omega$ plus anharmonic corrections, where $n = 0, 1, 2, \dots$. Mulliken studied the two molecules $\text{B}^{10}\text{O}^{16}$ and $\text{B}^{11}\text{O}^{16}$. He found that a good fit to the emission spectra could be obtained only if zero-point energy were included, or in his words, "if one assumes that the true values of the vibrational quantum numbers are not n and n' but each $\frac{1}{2}$ unit greater ... It is then probable that the minimum vibrational energy of BO (and doubtless of other) molecules is $\frac{1}{2}$ quantum." It is worth noting that Mulliken reached this conclusion based on his spectroscopic data, before Heisenberg (1925) derived the zero-point energy of a harmonic oscillator from matrix mechanics.

XIII. SUMMARY AND REMARKS

Zero-point energy first appeared in Planck's "second theory" of blackbody radiation. The concept was quickly adopted by Einstein and Stern, who showed that it could be used to derive the Planck spectrum from largely classical considerations. They also showed that rotational zero-point energy might account for an observed decrease with temperature of the specific heat of molecular hydrogen. None of these ingenious theories turned out to be quite correct from a modern perspective.

Zero-point motion played no role in Einstein's epiphanic paper of 1917 in which he derived the Planck spectrum using his A and B coefficients. It was the great simplicity of Einstein's derivation, perhaps, that ended speculations about the role of zero-point energy in the blackbody problem. However, we have seen that zero-point energy of the electromagnetic field has something to do with the A coefficient for spontaneous emission, although it cannot be regarded as the sole "cause" of emission. Furthermore, this zero-point energy appears to be important in connection with the Einstein fluctuation formula, historically the first indicator of wave-particle dualism, and Bose-Einstein statistics.

We have described how zero-point energy appeared and

was used during the development of quantum theory. Although interest in the concept in connection with black-body theory declined after Einstein's 1917 paper, it was by no means abandoned. In particular, direct spectroscopic evidence for the reality of zero-point energy was provided by Mulliken in 1924, just months before it appeared so naturally in the quantum formalism established in 1925–1926. And during the 1930s, the concept of zero-point energy was applied to a variety of problems in physical chemistry.²⁷ Stern in 1913 had already employed zero-point energy in a calculation of the vapor pressure of solids,^{27,28} and according to Enz,²⁸ Stern tried for some time to “convert Pauli to the zero-point energy against which he had the gravest hesitations.”

For two decades after the invention of the quantum formalism, physicists for the most part seemed not to pay much attention to zero-point energy. This changed in the late 1940s with the measurement of the Lamb shift and its interpretation as an effect of the zero-point electromagnetic field. And around that time Casimir described several implications of zero-point electromagnetic radiation, including the “Casimir force” between two conducting plates. With the full development of quantum theory “all the consequences of Planck's second hypothesis were proved although the point of view became very different. It would seem that the careful, cautious, and conservative Max Planck had exhibited in some decisive questions about nature more prescience than many a so-called revolutionary physicist.”²⁷

The idea of the vacuum as a quantum state with zero-point energy and fluctuations of physical consequence is now a commonplace, with implications ranging from optical communications to quantum chromodynamics to inflationary models of the Universe. From relatively humble origins in the second decade of this century and after long periods of neglect, the concept of zero-point energy has become firmly ingrained in the worldview of contemporary physicists.

ACKNOWLEDGMENTS

We thank a referee for noting some historical and technical imprecision in the original manuscript, and Professor S. G. Brush for kindly responding to our request for information concerning literature mentioned by the referee.

APPENDIX A: OSCILLATOR EQUATION AND ABSORPTION RATE

The Newton equation of motion for a particle of mass m and charge e , acted upon by an elastic restoring force $-m\omega_0^2 z$ and an external electric field $E_z(t)$, is

$$\frac{d^2 z}{dt^2} + \omega_0^2 z = \frac{e}{m} E_z(t) + \frac{e}{m} E_{RR}(t). \quad (A1)$$

For simplicity, and to follow Planck, Einstein, and Hopf, we assume the particle is constrained to one-dimensional motion.

The field $E_{RR}(t)$ in (A1) is the field of radiation reaction, i.e., the electric field produced by the charged particle at the position of the particle. In other words, it is the electric field that the charge exerts on itself. For our purposes here a simplified derivation and expression for this field will suffice.

We recall first the expression (8) for the rate at which an

accelerating charge radiates electromagnetic energy. The energy radiated in the time interval from t_1 to t_2 is

$$\begin{aligned} W_{EM}(t_2, t_1) &= \frac{2e^2}{3c^3} \int_{t_1}^{t_2} \left(\frac{d^2 z}{dt^2} \right)^2 dt \\ &= \frac{2e^2}{3c^3} \left(\frac{d^2 z(t)}{dt^2} \frac{dz(t)}{dt} \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{d^3 z}{dt^3} \frac{dz}{dt} dt \right), \end{aligned} \quad (A2)$$

where the second equality follows from an integration by parts. We assume the motion of the charge is periodic and choose $t_2 - t_1$ to be an integral number of periods, in which case

$$W_{EM}(t_2, t_1) = -\frac{2e^2}{3c^3} \int_{t_1}^{t_2} \frac{d^3 z}{dt^3} \frac{dz}{dt} dt. \quad (A3)$$

The change in energy of the charge, $-W_{EM}$, is attributed to the force $eE_{RR}(t)$ of radiation reaction:

$$\begin{aligned} -W_{EM}(t_2, t_1) &= \frac{2e^2}{3c^3} \int_{t_1}^{t_2} \frac{d^3 z}{dt^3} \frac{dz}{dt} dt \\ &= \int_{t_1}^{t_2} eE_{RR}(t) \frac{dz}{dt} dt \end{aligned} \quad (A4)$$

or

$$E_{RR}(t) = \frac{2e}{3c^3} \frac{d^3 z}{dt^3}. \quad (A5)$$

Although this expression for the radiation reaction field was derived under the assumption of periodic motion, it actually holds more generally. When it is used in (A1), we obtain the equation (41) used by Planck, Einstein, Hopf, and others.

For the case of a monochromatic applied field $E_z(t) = E_{z0} \cos(\omega t + \theta_\omega)$, Eq. (41) has the solution

$$z(t) = \text{Re} \left(\frac{(-e/m) E_{z0} e^{-i(\omega t + \theta_\omega)}}{\omega^2 - \omega_0^2 + i\gamma\omega^3} \right), \quad (A6)$$

so that the rate at which the oscillator absorbs energy from the field is found after some simple algebra to be

$$\dot{W}_A = e\dot{z}(t)E_z(t) \rightarrow \frac{e^2}{2m} \frac{\gamma\omega^4 E_{z0}^2}{(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^6}, \quad (A7)$$

where we have taken an average over the oscillations of the field, replacing $\cos^2(\omega t + \theta_\omega)$ by $1/2$ and $\sin(\omega t + \theta_\omega) \cos(\omega t + \theta_\omega)$ by 0 .

Now suppose the applied field has a broad distribution of frequencies, with energy density in the interval $[\omega, \omega + d\omega]$ given by $\rho(\omega)d\omega = E_{z0}^2/8\pi$. In this case, (A7) is replaced by

$$\dot{W}_A = \frac{4\pi e^2}{m} \int_0^\infty \frac{\omega^4 \rho(\omega) d\omega}{(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^6}. \quad (A8)$$

The time $\gamma = 2e^2/3mc^3 = 6.3 \times 10^{-24}$ s is so short that for natural oscillation frequencies ω_0 of interest $\gamma\omega_0 \ll 1$. Furthermore, $\rho(\omega)$ may be assumed to be flat compared with the sharply peaked function

$$\frac{\omega^4}{(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^6} \approx \frac{\omega_0^4}{4\omega_0^2(\omega - \omega_0)^2 + \gamma^2\omega_0^6}, \quad (A9)$$

in the integrand of (A8), so that

$$\begin{aligned} \dot{W}_A &\cong \frac{\pi e^2 \gamma \omega_0^2 \rho(\omega_0)}{m} \int_0^\infty \frac{d\omega}{(\omega - \omega_0)^2 + \gamma^2 \omega_0^2 / 4} \\ &\cong \frac{\pi e^2 \gamma \omega_0^2 \rho(\omega_0)}{m} \left(\frac{2\pi}{\gamma \omega_0^2} \right) \\ &\cong \frac{2\pi^2 e^2}{m} \rho(\omega_0) = \frac{\pi e^2}{m} \rho(\nu_0) \rightarrow \frac{\pi e^2}{3m} \rho(\nu_0). \end{aligned} \quad (\text{A10})$$

In the last step, we have replaced $\rho(\omega)$ by $\rho(\nu_0)/3$, where now the spectral energy density is defined by $\rho(\omega)d\omega = (E_{x\omega}^2 + E_{y\omega}^2 + E_{z\omega}^2)/8\pi = 3E_{z\omega}^2/8\pi$ for (isotropic and unpolarized) thermal radiation. We have thus arrived at Eq. (7) for the energy absorption rate.

By replacing e^2/m by $e^2 f/m$ in Eq. (7), where f is the oscillator strength of an atomic transition of frequency ω_0 , we obtain the energy absorption rate given by quantum mechanics up to second order in perturbation theory.²⁹

APPENDIX B: FORCE ON AN ATOM IN A THERMAL FIELD

We shall follow Einstein's derivation of the force acting on an atom moving with velocity v in a thermal field. A classical derivation can be given along similar lines for the Einstein-Hopf force (42) acting on a classical dipole oscillator. Since the result differs from (93) only by simple multiplicative factors, we will not go through the classical derivation here.

The field energy density in the frequency interval $[\omega, \omega + d\omega]$ and within the solid angle $d\Omega$ is $\rho(\omega)d\omega d\Omega/4\pi$, where $\rho(\omega)$ is independent of direction since thermal radiation is isotropic in the laboratory frame. Consider radiation propagating in a direction θ with respect to the axis defined by the atom's velocity. The frequency of radiation in the atom's frame is Doppler shifted to

$$\omega' \cong \omega [1 - (v/c) \cos \theta] \quad (v/c \ll 1). \quad (\text{B1})$$

The radiation appears to the atom to be directed at an angle θ' given by the aberration formula

$$\cos \theta' \cong \cos \theta - (v/c) \sin^2 \theta. \quad (\text{B2})$$

The field energy density $\rho'(\omega', \theta') d\omega' d\Omega'/4\pi$ in the frame of the moving atom can be obtained straightforwardly from the well-known transformation properties of the electric and magnetic fields under Lorentz transformations of the coordinates. We simply write the result:

$$\rho'(\omega', \theta') d\omega' d\Omega' \cong [1 - 2(v/c) \cos \theta] \rho(\omega) d\omega d\Omega \quad (\text{B3})$$

or

$$\begin{aligned} \rho'(\omega', \theta') &\cong \left(1 - 2 \frac{v}{c} \cos \theta \right) \rho(\omega) \frac{d\omega}{d\omega'} \frac{d(\cos \theta)}{d(\cos \theta')} \\ &\cong [1 - (3v/c) \cos \theta'] \rho(\omega), \end{aligned} \quad (\text{B4})$$

where we have used (B1) and (B2) and continue to assume $v/c \ll 1$. From (B1) it also follows that

$$\begin{aligned} \rho(\omega) &\cong \rho[\omega' + (v\omega'/c) \cos \theta'] \\ &\cong \rho(\omega') + \frac{d\rho(\omega')}{d\omega} \left(\frac{v}{c} \right) \omega' \cos \theta', \end{aligned} \quad (\text{B5})$$

so that (B4) becomes

$$\begin{aligned} \rho'(\omega', \theta') &\cong [1 - 3(v/c) \cos \theta'] \\ &\times \left[\rho(\omega') + \frac{d\rho(\omega')}{d\omega} \left(\frac{v}{c} \right) \omega' \cos \theta' \right]. \end{aligned} \quad (\text{B6})$$

Radiation in the solid angle $d\Omega'$ induces in the moving atom an average number

$$n_2 = B_{12} N_2 \rho'(\omega', \theta') d\Omega' / 4\pi, \quad (\text{B7})$$

of stimulated emission transitions per unit time, and a number

$$n_1 = B_{12} N_1 \rho'(\omega', \theta') d\Omega' / 4\pi, \quad (\text{B8})$$

of absorption transitions. The net momentum per unit time imparted to the atom due to stimulated emission and absorption of photons of momentum $\hbar\omega'/c$ is thus

$$F = \frac{dp}{dt} = (n_1 - n_2) \frac{\hbar\omega'}{c} \cos \theta', \quad (\text{B9})$$

since absorption causes the atom to recoil in the same direction as the field propagation, whereas, from conservation of linear momentum stimulated emission causes recoil in the opposite direction. From (B7)–(B9),

$$F = (\hbar\omega'/c) (B_{12}/4\pi) (N_1 - N_2) \rho'(\omega', \theta') \cos \theta' d\Omega'. \quad (\text{B10})$$

Note that spontaneous emission adds no net momentum on average to the atom since it is equally likely in all directions.

We now add up the forces associated with all directions of propagation of radiation of frequency ω :

$$\begin{aligned} F &= \frac{\hbar\omega'}{c} \frac{B_{12}}{4\pi} (N_1 - N_2) \int_0^{2\pi} d\phi' \int_0^\pi d\theta' \\ &\times \sin \theta' \rho'(\omega', \theta') \cos \theta' \\ &= \frac{\hbar\omega'}{2c} \frac{B_{12}}{4\pi} (N_1 - N_2) \int_0^\pi d\theta' \sin \theta' \cos \theta' \rho'(\omega', \theta') \\ &\cong - \left(\frac{\hbar\omega}{c^2} \right) (N_1 - N_2) B_{12} \left(\rho(\omega) - \frac{\omega}{3} \frac{d\rho(\omega)}{d\omega} \right) v, \end{aligned} \quad (\text{B11})$$

where to lowest order in v/c we have dropped all primes. This is Einstein's equation (93) for the force on an atom moving in a thermal field.

²⁹This paper is based on the introductory chapter of *The Quantum Vacuum*, by P. W. Milonni and M. -L. Shih, to be published by Academic in 1992. The authors are grateful to Robert Kaplan, Senior Editor, for permission to submit this material for publication in *The American Journal of Physics*.

¹M. Planck, "Über die Begründung des Gesetzes der schwarzen Strahlung," *Ann. Phys. (Leipzig)* **37**, 642–656 (1912).

²For references to early work on the blackbody problem, see T. S. Kuhn, *Blackbody Theory and the Quantum Discontinuity*, 1894–1912 (Oxford U. P., New York, 1978).

³A plane wave exerts a pressure $2u$ on a reflecting surface on which it is normally incident. For plane waves propagating with equal intensities in both directions normal to the surface, this is reduced to u , or $u/3$ if the radiation is isotropic.

⁴For a discussion of the classical oscillator model of an atom, often called the Lorentz model, see P. W. Milonni and J. H. Eberly, *Lasers* (Wiley, New York, 1988), Chaps. 2 and 3.

⁵Lord Rayleigh, "Remarks upon the law of complete radiation," *Philos. Mag.* **49**, 539–540 (1900). It is perhaps worth noting that near the end of his note Rayleigh modified his formula of the form (12) by including an exponential factor $e^{-c\nu/\lambda T}$. This modification, based on the earlier Wien distribution, of course avoided the ultraviolet catastrophe, although it is

not clear that this was the motivation behind Rayleigh's inclusion of the exponential factor. Kuhn (Ref. 2, p. 145) remarks that "Rayleigh's 1900 note is both cryptic and incomplete."

⁶This is sometimes called Jeans' theorem.

⁷See A. Pais, "Subtle is the Lord" ... *The Science and the Life of Albert Einstein* (Oxford U. P., New York, 1982).

⁸See, for instance, the remarks of Einstein, in *Albert Einstein: Philosopher-Scientist*, edited by P. A. Schillp (Tudor, New York, 1949), p. 45.

⁹Reference 2, pp. 117–118.

¹⁰A. Einstein and O. Stern, "Einige Argumente für die Annahme einer molekularen Agitation beim absoluten Nullpunkt," *Ann. Phys. (Leipzig)* **40**, 551–560 (1913). An English translation with comments is given by S. Bergia, P. Lugli, and N. Zamboni, "Zero-point energy, Planck's law and the prehistory of stochastic electrodynamics. Part 2. Einstein and Stern's paper of 1913," *Ann. Fond. Louis de Broglie* **5**, 39–62 (1980).

¹¹A. Einstein, "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt," *Ann. Phys. (Leipzig)* **17**, 132–148 (1905).

¹²A. Einstein, "Zur Theorie der Lichterzeugung und Lichtabsorption," *Ann. Phys. (Leipzig)* **20**, 199–207 (1906).

¹³A. Einstein and L. Hopf, "Statistische Untersuchung der Bewegung eines Resonators in einem Strahlungsfeld," *Ann. Phys. (Leipzig)* **33**, 1105–1115 (1910). An English translation with comments is given by S. Bergia, P. Lugli, and N. Zamboni, "Zero-point energy, Planck's law and the prehistory of stochastic electrodynamics. Part 1. Einstein and Hopf's paper of 1910," *Ann. Fond. Louis de Broglie* **4**, 295–318 (1979).

¹⁴A. Einstein and L. Hopf, "Über einen Satz der Wahrscheinlichkeitsrechnung und seine Anwendung in der Strahlungstheorie," *Ann. Phys. (Leipzig)* **33**, 1096–1104 (1910).

¹⁵P. W. Milonni, "Quantum mechanics of the Einstein–Hopf model," *Am. J. Phys.* **49**, 177–184 (1981).

¹⁶See Ref. 8, p. 43.

¹⁷W. Nernst, "Über einen Versuch von quantentheoretischen Betrachtungen zur Annahme stetiger Energieänderungen zurückzukehren,"

Ver. Dsch. Phys. Gen. **18**, 83 (1916).

¹⁸A. Einstein, "Zur gegenwärtigen Stand des Strahlungsproblems," *Phys. Z.* **10**, 185–193 (1909).

¹⁹See, for instance, P. W. Milonni, "Wave-particle duality of light: A current perspective," in *The Wave-Particle Dualism*, edited by S. Diner, D. Fargue, G. Lochak, and F. Selleri (Reidel, Dordrecht, 1984), pp. 27–67.

²⁰A. Einstein, "Zur Quantentheorie der Strahlung," *Phys. Z.* **18**, 121–128 (1917).

²¹V. F. Weisskopf, "The development of field theory in the last 50 years," *Phys. Today* **34**(11), 69–85 (1981). See also Ref. 22 and V. L. Ginzburg, "The Nature of Spontaneous Radiation," *Usp. Fiz. Nauk* **40**, 687–698 (1983) [*Sov. Phys. Usp.* **26**, 713–719 (1983)].

²²See P. W. Milonni, "Why spontaneous emission?," *Am. J. Phys.* **52**, 340–343 (1984).

²³A. Einstein, "Die Plancksche Theorie der Strahlung und die Theorie der spezifischen Wärme," *Ann. Phys. (Leipzig)* **22**, 180–190 (1907).

²⁴A. Eucken, "Die Molekularwärme des Wasserstoffs bei tiefen Temperaturen," *Sitzungsber. Preuss. Akad. Wiss.* **1912**, 141–151 (February, 1912). See also Ref. 27.

²⁵P. Debye, "Interferenz von Röntgenstrahlen und Wärmebewegung," *Ann. Phys. (Leipzig)* **43**, 49–95 (1914).

²⁶R. S. Mulliken, "The band spectrum of boron monoxide," *Nature* **114**, 349–350 (1924).

²⁷J. Mehra and H. Rechenberg, "Planck's half-quanta: History of the concept of zero-point energy in the general development of quantum theory," report of the Center for Particle Theory, University of Texas at Austin, 1970.

²⁸C. P. Enz, "Is the zero-point energy real?," in *Physical Reality and Mathematical Description*, edited by C. P. Enz and J. Mehra (Reidel, Dordrecht, 1974), pp. 124–132.

²⁹See, for instance, M. Cray, M. -L. Shih, and P. W. Milonni, "Stimulated emission, absorption, and interference," *Am. J. Phys.* **50**, 1016–1021 (1982).

Simply no hidden variables

Thomas F. Jordan

Physics Department, University of Minnesota, Duluth, Minnesota 55812

E. C. G. Sudarshan

Center for Particle Theory, University of Texas, Austin, Texas 78712

(Received 18 June 1990; accepted for publication 31 January 1991)

It is shown very simply that for any two observables represented by Hermitian operators that do not commute, there is a state for which there is no joint probability distribution for the two observables.

I. INTRODUCTION

Does quantum mechanics not allow hidden variables? If so, it is not easy to see why from the classic proofs.^{1–3} They are too complicated. They assume as little as possible and prove as much as possible. To do so, they avoid full use of quantum mechanics^{1,2} and/or use unfamiliar language involving a lattice of propositions² or partial algebra.³ They left an impression that questions about hidden variables are only for specialists.

That has changed. Hidden variables are used to derive Bell inequalities.⁴ We see the absence of these hidden variables as a basic property of nature which quantum mechanics accommodates and describes. It was tested when experi-

ments⁵ showed Bell inequalities are false. Still, arguments for these hidden variables, involving separated subsystems, are the most difficult to dismiss.

Recently, Sudarshan and Rothman⁶ pointed out that calculations of correlations in quantum mechanics can be done in a way that is similar in structure to derivations of Bell inequalities. It highlights the key difference. In derivations of Bell inequalities, the distributions used to calculate correlations are supposed to be actual nonnegative probabilities. They are joint probability distributions for observables that in quantum mechanics are represented by operators that do not commute.⁷ Fine⁸ showed that existence of these joint probability distributions is equivalent to the Bell inequalities and equivalent to the assumptions about hid-