painless short course, I think, for those crossing the relativistic Rubicon for the first or second time.

Originally I hoped to cover more ground, including chapters on the vacuum in quantum chromodynamics and cosmology. However, I gradually realized that in a book of the intended size I could not do this without sacrificing the rather considerable detail I wanted for various QED effects, particularly van der Waals and Casimir interactions. Since these things are not treated in any sort of detail in other books I know of, I decided to omit the somewhat more speculative aspects of vacuum theory and to focus on quantum electrodynamics.

Writing this book has given me much pleasure. I wish to thank especially Mei-Li Shih and Gordon W. F. Drake. Mei-Li not only was relentless in her encouragement but also checked many of the calculations and made good suggestions. In my opinion, but not hers, she is an author of this book. Professor Drake served as a reviewer for Academic Press and, by suggesting changes and identifying errors and non sequiturs, helped greatly to improve the quality of the manuscript.

Discussions relating to this book with Richard J. Cook, Joseph H. Eberly, Walter T. Grandy, Jr., Michael Lieber, David Nesbitt, M.D., Edwin A. Power, David H. Sharp, and Larry Spruch were enlightening and often encouraging. Professor Eberly supervised the Ph.D thesis work from which this book gradually evolved.

I think it is also appropriate to thank Gabriel Barton, Timothy H. Boyer, Trevor W. Marshall, and again, Edwin Power; their work has been partly responsible for my continued fascination with the subject.

Needless to say, H. B. G. Casimir has given us a lot to think about, and I thank him for his kind replies to my letters and for permission to quote from them.

Finally I wish to acknowledge the inestimable help I have received from the editors at Academic Press. I hope this book lives up to their high standards and expectations.

Peter W. Milonni
Los Alamos, New Mexico

# Chapter 1

# Zero-Point Energy in Early Quantum Theory

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

## 1.1 Introduction

The importance of the blackbody problem in the development of quantum theory is recognized by every serious student of modern physics. What is not so widely known is that blackbody theory led also to the concept of zero-point energy, which was later to appear naturally in the mathematics of quantum theory. The relation of this energy to early premonitions of wave-particle dualism is similarly not widely appreciated. This chapter 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.

## 1.2 The Blackbody Problem

In 1860 Kirchhoff derived a general relation between the radiative and absorptive strengths of a body held at a fixed temperature $T$. According to Kirchhoff's law the ratio of the radiative strength to the absorption coefficient for radiation of wavelength $\lambda$ 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 every

wavelength, corresponding to total absorption. Thus $F(\lambda, T)$ characterizes the radiative strength at wavelength $\lambda$ 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.[1] 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

$$dQ = dU + PdV = d(uV) + \frac{1}{3}udV = Vdu + \frac{4}{3}udV \tag{1.1}$$

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

$$dQ = V\frac{du}{dT}dT + \frac{4}{3}udV. \tag{1.2}$$

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

$$dS = \frac{1}{T}dQ = \frac{V}{T}\frac{du}{dT}dT + \frac{4}{3}\frac{u}{T}dV, \tag{1.3}$$

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

$$\frac{\partial S}{\partial T} = \frac{V}{T}\frac{du}{dT}, \quad \frac{\partial S}{\partial V} = \frac{4}{3}\frac{u}{T} \tag{1.4}$$

and

$$\frac{\partial^2 S}{\partial T \partial V} = \frac{\partial^2 S}{\partial V \partial T} = \frac{1}{T}\frac{du}{dT} = \frac{4}{3}\frac{d}{dT}\left(\frac{u}{T}\right), \tag{1.5}$$

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

$$u = bT^4 \quad \text{(Stefan–Boltzmann law)}, \tag{1.6}$$

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

[1] A plane wave exerts a pressure $2u$ on a reflecting surface on which it is normally incident. (See, for instance, W. K. H. Panofsky and M. Phillips, *Classical Electricity and Magnetism* (Addison-Wesley, Reading, Mass., 1962), p. 193.) 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.



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. 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)

$$\dot{W}_{\mathrm{A}} = \frac{\pi e^2}{3m}\rho(\nu_o), \tag{1.7}$$

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

$$\dot{W}_{\mathrm{EM}} = \frac{2e^2a^2}{3c^3}, \tag{1.8}$$

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

$$\dot{W}_{\mathrm{EM}} = \left(\frac{32\pi^4 e^2 \nu_o^4}{3c^3}\right)x^2, \tag{1.9}$$

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 $\frac{1}{2}m\omega_o^2 x^2$ 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 (1.7) should equal the emission rate (1.9). Thus

$$\rho(\nu_o) = \frac{8\pi\nu_o^2}{c^3}(m\omega_o^2 x^2) = \frac{8\pi\nu_o^2}{c^3}U, \tag{1.10}$$

or more generally

$$\rho(\nu) = \frac{8\pi\nu^2}{c^3}U \tag{1.11}$$

for a blackbody, which absorbs at all frequencies $\nu$. 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

$$\rho(\nu) = \left(\frac{8\pi\nu^2}{c^3}\right)kT \quad \text{(Rayleigh–Jeans distribution)}. \tag{1.12}$$

The total electromagnetic energy density

$$u = \int_0^\infty \rho(\nu) d\nu \tag{1.13}$$

violates the Stefan–Boltzmann law. Furthermore the Rayleigh–Jeans law suffers from the ultraviolet catastrophe: $u$ diverges when (1.12) is used for $\rho(\nu)$.

Equation (1.11) was derived by Planck and, as we shall see, played a very important role in his work on the blackbody problem.

Equation (1.12) for the spectral energy density of blackbody radiation was first deduced in a less explicit form by Rayleigh in 1900.[2] 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 $\nu$ is basically just a linear harmonic oscillator (see Chapter 2) that, according to the classical equipartition theorem, has an average energy $kT$ at thermal equilibrium. 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, independent 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

$$\rho(\nu) = \nu^3 \phi_1(\nu/T) \quad \text{(Wien displacement law)}, \tag{1.14}$$

or, in terms of wavelength,

$$\rho(\lambda) = \rho(\nu)\left|\frac{d\nu}{d\lambda}\right| = \lambda^{-5}\phi_2(\lambda T) \quad \text{(Wien displacement law)}, \tag{1.15}$$

where $\phi_1$ and $\phi_2$ are undetermined functions. The Rayleigh–Jeans distribution obviously obeys Wien's "displacement law" (1.14).

[2] Motivated by Wien's work, Rayleigh also allowed for the possiblility that a factor $e^{-(\text{const})\nu/T}$ should be included, thus avoiding the ultraviolet catastrophe.



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

$$\rho(\lambda) = \alpha\lambda^{-5}e^{-\beta/\lambda T} \quad \text{(Wien distribution)}, \tag{1.16}$$

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

Wien's arguments for (1.16) 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 (1.16) 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.

## 1.3 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 the *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 $T$ :

$$\frac{\partial^2 S}{\partial U^2} = -\frac{A}{U}, \tag{1.17}$$

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

$$U = Be^{-1/AT}, \tag{1.18}$$

where $B$ is another constant that, like $A$, may depend on the frequency of a given radiator. This result, together with (1.11), yields the radiation spectral energy density

$$\rho(\nu) = f(\nu)e^{-1/AT}, \tag{1.19}$$

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

$$\rho(\nu) = C\nu^3 e^{-D\nu/T} \quad (C, D \text{ constants}) \tag{1.20}$$

or

$$\rho(\lambda) = \alpha\lambda^{-5} e^{-\beta/\lambda T} \quad (\alpha, \beta \text{ constants}), \tag{1.21}$$

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 $\mu$m, 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. In particular, the data indicated that $\rho(\nu)$ was proportional to the temperature $T$ for small $\nu$ and large $T$. Planck found a formula with that behavior at small $\nu$ and which approximated the Wien distribution for large $\nu$.

In a paper delivered at a meeting on 19 October Planck presented his formula and provided some justification for it.[3] For small $\nu$ and large $T$, the experimental result $\rho(\nu) \propto T$ and equation (1.11) 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 (1.17) leads to the Wien distribution, which has the correct form for large $\nu$ and small $T$. Planck proposed the interpolation

$$\frac{\partial^2 S}{\partial U^2} = \frac{-A}{U(B+U)} \quad (A, B \text{ constants}). \tag{1.22}$$

According to Planck, equation (1.22) "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 $\partial S/\partial U = 1/T$, equation (1.11), and the Wien displacement law, one obtains from (1.22) the spectral energy density

$$\rho(\lambda) = \frac{\alpha\lambda^{-5}}{e^{\beta/\lambda T} - 1} \quad (\alpha, \beta \text{ constants}). \tag{1.23}$$

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

[3] See the books by Kuhn and Pais and the articles by Klein cited at the end of the chapter.



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 glibly summarized as follows. Consider $N$ radiators of frequency $\nu$ and total energy $U_N = NU = P\epsilon$, where $P$ is a large integer and $\epsilon$ 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 $(2\epsilon, 0)$, $(\epsilon, \epsilon)$, and $(0, 2\epsilon)$ *if the energy elements are assumed to be indistinguishable.* Under this assumption we have, in general,

$$W_N = \frac{(N-1+P)!}{P!(N-1)!}, \tag{1.24}$$

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 $N, P >> 1$,

$$\begin{aligned} S &= \frac{k}{N} \log \frac{(N-1+P)!}{P!(N-1)!} \\ &\cong k\left[(1+\frac{P}{N})\log(1+\frac{P}{N}) - \frac{P}{N}\log\frac{P}{N}\right] \\ &= k\left[(1+\frac{U}{\epsilon})\log(1+\frac{U}{\epsilon}) - \frac{U}{\epsilon}\log\frac{U}{\epsilon}\right]. \end{aligned} \tag{1.25}$$

Thus

$$\frac{\partial S}{\partial U} = \frac{1}{T} = \frac{k}{\epsilon}\log(1+\frac{\epsilon}{U}) \tag{1.26}$$

or

$$U = \frac{\epsilon}{e^{\epsilon/kT} - 1} \tag{1.27}$$

for the average energy of each radiator. The excellent agreement between (1.23) and experiment, together with equation (1.11), suggests that $\epsilon$ is inversely proportional to the wavelength, or directly proportional to the frequency of the oscillator:

$$\epsilon = h\nu. \tag{1.28}$$

Then

$$U = \frac{h\nu}{e^{h\nu/kT} - 1} \tag{1.29}$$

and (1.11) implies

$$\rho(\nu) = \frac{8\pi h\nu^3/c^3}{e^{h\nu/kT} - 1} \quad \text{(Planck spectrum)} \tag{1.30}$$

for the spectral energy density of thermal radiation.

The expression (1.25) for $S$ satisfies equation (1.22) with $A = k$ and $B = \epsilon$. Once (1.25) is obtained, therefore, one is led to the form (1.23) for the spectral energy density. The great success of (1.23) in fitting the experimental data led Planck to what he later called an "act of desperation" needed to derive (1.25).

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. His "energy elements" obeyed what would much later be recognized as Bose–Einstein statistics.

Another revolutionary (nonclassical) aspect of Planck's calculation, of course, is the physical significance it attaches to the "energy elements" of size $\epsilon$, and the relation (1.28) between $\epsilon$ and the frequency $\nu$ of a *material* oscillator. Boltzmann had also employed "energy elements" in his counting of complexions, but in his calculations $\epsilon$ 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 $\epsilon \to 0$ in equation (1.26), however, then $\partial S/\partial U \to k/U$ and $\partial^2 S/\partial U^2 \to -k/U^2$, which leads to the Rayleigh–Jeans distribution. In Planck's derivation of his spectrum, therefore, the quantization of energy was absolutely essential.

This is the traditional view of Planck's innovation. It should be noted, however, that Kuhn (1978) has concluded that Planck did not in 1900 introduce any physical quantization of either radiation or material radiators. He argues that Planck's radiators were 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."

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 (1913), is worth mentioning. Consider the classical limit $kT >> h\nu$ of the expression (1.29) for the average energy of an



oscillator in thermal equilibrium with radiation:

$$U = \frac{h\nu}{e^{h\nu/kT}-1} \cong \frac{h\nu}{1+\frac{h\nu}{kT}+\frac{1}{2}(\frac{h\nu}{kT})^2-1} = \frac{kT}{1+\frac{1}{2}\frac{h\nu}{kT}} \cong kT - \frac{1}{2}h\nu. \quad (1.31)$$

Thus $U$ contains a first-order temperature-independent correction to $kT$, the energy predicted by the equipartition theorem in the classical limit. But

$$U + \frac{1}{2}h\nu = \frac{h\nu}{e^{h\nu/kT}-1} + \frac{1}{2}h\nu, \quad (1.32)$$

*which includes the zero-point energy* $\frac{1}{2}h\nu$, does not have a first-order correction to $kT$ in the classical limit. In Planck's "second theory" $U$ was in fact replaced by $U + \frac{1}{2}h\nu$.

## 1.4 Planck's Zero-Point Energy

It was mentioned earlier 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." 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 $h\nu$. Let $P_n$ be the probability that it has energy between $(n-1)h\nu$ and $nh\nu$. When, as a result of absorption of radiation, its energy reaches $nh\nu$, 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 $P_2 = P_1(1-p)$, $P_3 = P_2(1-p) = P_1(1-p)^2, \ldots,$ $P_n = P_1(1-p)^{n-1}$, and

$$\sum_{n=1}^{\infty} P_n = 1 = \sum_{n=1}^{\infty} P_1(1-p)^{n-1} = P_1/p \quad (1.33)$$

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

$$S = -k\sum_{n=1}^{\infty} P_n \log P_n = -k\sum_{n=1}^{\infty} p(1-p)^{n-1}\log[p(1-p)^{n-1}]$$

$$= -k\left[\frac{1}{p}\log p + (\frac{1}{p}-1)\log(\frac{1}{p}-1)\right]. \quad (1.34)$$

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

$$U = \sum_{n=1}^{\infty}(n-\frac{1}{2})h\nu P_n = h\nu\sum_{n=1}^{\infty}(n-\frac{1}{2})p(1-p)^{n-1} = (\frac{1}{p}-\frac{1}{2})h\nu \quad (1.35)$$

or $1/p = U/h\nu + \frac{1}{2}$. From (1.34), therefore,

$$S = k\left[(\frac{U}{h\nu}+\frac{1}{2})\log(\frac{U}{h\nu}+\frac{1}{2}) - (\frac{U}{h\nu}-\frac{1}{2})\log(\frac{U}{h\nu}-\frac{1}{2})\right]. \quad (1.36)$$

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

$$U = \frac{1}{2}h\nu\frac{e^{h\nu/kT}+1}{e^{h\nu/kT}-1} = \frac{h\nu}{e^{h\nu/kT}-1} + \frac{1}{2}h\nu. \quad (1.37)$$

This implies that $U \neq 0$ when $T \to 0$: when $T \to 0$, $U \to \frac{1}{2}h\nu$. *Planck's equation (1.37) marked the birth of the concept of zero-point energy.*

To derive $\rho(\nu)$ Planck could not resort to equation (1.11), 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 (1.35), $U = [C\rho(\nu)+\frac{1}{2}]h\nu$ or

$$\rho(\nu) = \frac{1}{C}\frac{1}{e^{h\nu/kT}-1}. \quad (1.38)$$

To determine $C$ Planck appeals to the classical limit, where the Rayleigh–Jeans law should apply: for $kT >> h\nu$, $\rho(\nu)$ should reduce to (1.12), which requires that $1/C = 8\pi h\nu^3/c^3$ and therefore that

$$\rho(\nu) = \frac{8\pi h\nu^3/c^3}{e^{h\nu/kT}-1}. \quad (1.39)$$

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.

It is also noteworthy that in Planck's second theory the material oscillators have zero-point energy but the electromagnetic field does not: $\rho(\nu) \to 0$ for $T \to 0$. Had Planck simply used equation (1.11) to relate $\rho(\nu)$ and $U$, he would have obtained from (1.37) the spectral energy density

$$\rho'(\nu) = \rho(\nu) + 4\pi h\nu^3/c^3 = \frac{8\pi h\nu^3/c^3}{e^{h\nu/kT}-1} + 4\pi h\nu^3/c^3, \quad (1.40)$$

which, as we will see later, turns out to be the correct spectrum from the standpoint of modern quantum electrodynamics. The zero-point energy appearing in Planck's expression (1.37) is also perfectly correct according to modern theory, even though Planck's route to it is not.

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.

## 1.5 The Einstein–Hopf Model

"Concerning a Heuristic Point of View Toward the Emission and Transformation of Light," Einstein (1905) deduced that radiation satisfying the Wien distribution "behaves thermodynamically as though it consisted of a number of independent energy quanta of magnitude $[h\nu]$." 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 $h\nu$." These were the beginnings of the photon concept.

Einstein struggled with the blackbody problem for more than ten years after he introduced his heuristic viewpoint concerning energy quanta of radiation. In one important paper Einstein and Hopf (1910b) studied a simple model for the thermal equilibrium between oscillating dipoles and electromagnetic radiation. Imagine each dipole to consist of a particle of mass $m$ and charge $e$, bound by an elastic restoring force to a mass $M(>> m)$ of opposite charge. The equation of motion for a linear dipole oscillator is then (see Appendix A)

$$\frac{d^2z}{dt^2} + \omega_o^2 z - \tau\dddot{z} = \frac{e}{m}E_z(t), \quad (1.41)$$

where $\omega_o(= 2\pi\nu_o)$ is the natural oscillation frequency, $E_z(t)$ is the $z$-component of the external electric field acting on the particle, $\tau\dddot{z}$ is the

radiation reaction term, and $\tau = 2e^2/3mc^3$. The two oppositely charged particles define an electric dipole moment $ez(t)$. Implicit in equation (1.41) 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 (1.41) is essentially the same equation used earlier by Planck to derive equation (1.11) (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. Due to 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 $v/c << 1$, Einstein and Hopf showed that the retarding force due to motion through a thermal field of spectral energy density $\rho(\omega_o)$ is

$$F = -Rv, \tag{1.42}$$

where

$$R = \frac{4\pi^2 e^2}{5mc^2}\left[\rho(\omega_o) - \frac{\omega_o}{3}\frac{d\rho}{d\omega_o}\right] \tag{1.43}$$

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 $\delta t$ its momentum is

$$Mv(t+\delta t) = Mv(t) + \Delta - Rv(t)\delta t, \tag{1.44}$$

where $\Delta$ is the impulse imparted to the dipole in the time interval $\delta t$ as a result of recoil associated with emission and absorption of radiation. Then

$$M^2v^2(t+\delta t) - M^2v^2(t) = \Delta^2 - 2MRv^2(t)\delta t + (2M - R\delta t)v(t)\Delta \tag{1.45}$$

when $\delta t$ is taken to be small enough (or $M$ large enough) that terms quadratic in $\delta t$ are negligible. Now take the equilibrium ensemble average of both sides of (1.45):

$$2M[\langle\tfrac{1}{2}Mv^2(t+\delta t)\rangle - \langle\tfrac{1}{2}Mv^2(t)\rangle] = 0 = \langle\Delta^2\rangle - 4R\delta t\langle\tfrac{1}{2}Mv^2(t)\rangle. \tag{1.46}$$

In writing this expression we have used the fact that $\langle v(t)\Delta\rangle = 0$, since $\Delta$ is equally likely to be positive or negative in the time interval from $t$ to $t + \delta t$. 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

$$(\delta t)^{-1}\langle\Delta^2\rangle = 2RkT. \tag{1.47}$$

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



The force on an electric dipole moment $e\hat{z}z(t)$ in the Einstein–Hopf model, where the dipole points in the $z$ direction and is free to move only along $x$, may for our purposes be taken to be $F_x = e\hat{x}z(t)\partial E_z(t)/\partial x$.[4] The impulse imparted to the particle during the time interval from $t = 0$ to $t = \delta t$ is thus

$$\Delta = e\int_0^{\delta t} dt z(t)\frac{\partial E_z(t)}{\partial x}\,. \tag{1.48}$$

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

$$\mathbf{E}(\mathbf{r},t) = i\sum_{\mathbf{k}\lambda}[A_{\mathbf{k}\lambda}e^{-i(\omega_k t+\theta_{\mathbf{k}\lambda})} - A^*_{\mathbf{k}\lambda}e^{i(\omega_k t+\theta_{\mathbf{k}\lambda})}]\mathbf{e}_{\mathbf{k}\lambda}, \tag{1.49}$$

where $\mathbf{e}_{\mathbf{k}\lambda}$ is a unit polarization vector for a plane wave with wave vector $\mathbf{k}$ and linear polarization index $\lambda$ $(= 1, 2)$. The steady-state solution of equation (1.41) is then

$$z(t) = \frac{ie}{m}\sum_{\mathbf{k}\lambda}[F_{\mathbf{k}\lambda}e^{-i(\omega_k t+\theta_{\mathbf{k}\lambda})} - F^*_{\mathbf{k}\lambda}e^{i(\omega_k t+\theta_{\mathbf{k}\lambda})}], \tag{1.50}$$

where the origin of coordinates has been chosen to be at the position of the dipole and $F_{\mathbf{k}\lambda} = -A_{\mathbf{k}\lambda}e_{\mathbf{k}\lambda z}[\omega_k^2 - \omega_o^2 + i\tau\omega_k^3]^{-1}$, where $e_{\mathbf{k}\lambda z}$ is the $z$-component of $\mathbf{e}_{\mathbf{k}\lambda}$. In a separate paper Einstein and Hopf (1910a) show that $E_z(t)$ and $\partial E_z/\partial x$ must be treated as independent random variables in the time integral (1.48). It then follows from (1.48)–(1.50) by straightforward manipulations that $\langle\Delta\rangle = 0$ and

$$(\delta t)^{-1}\langle\Delta^2\rangle = \left(\frac{4\pi^4c^4\tau}{5\omega_o^2}\right)\rho^2(\omega_o), \tag{1.51}$$

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

Equation (1.47), together with (1.43) and (1.51), now gives a differential equation that must be satisfied by the spectral energy density of thermal radiation:

$$\rho(\omega) - \frac{\omega}{3}\frac{d\rho}{d\omega} = \left(\frac{\pi^2c^3}{3\omega^2kT}\right)\rho^2(\omega). \tag{1.52}$$

[4] See P. W. Milonni and M.-L. Shih, *Am. J. Phys.* **59**, 684 (1991).

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

$$\rho(\omega) = \frac{\omega^2 kT}{\pi^2 c^3}, \tag{1.53}$$

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.

## 1.6 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. First let us note that equation (1.11) allows us to write (1.52) in a form in which the average dipole energy $U$ appears explicitly:

$$\rho(\omega) - \frac{\omega}{3}\frac{d\rho}{d\omega} = \frac{1}{3kT}\rho(\omega)U. \tag{1.54}$$

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 (1.54) is then replaced by

$$\begin{aligned} \rho(\omega) - \frac{\omega}{3}\frac{d\rho}{d\omega} &= \frac{1}{3kT}\rho(\omega)U + \frac{\hbar\omega}{3kT}\rho(\omega) \\ &= \frac{\pi^2 c^3}{3\omega^2 kT}\rho^2(\omega) + \frac{\hbar\omega}{3kT}\rho(\omega) \\ &= \frac{\pi^2 c^3}{3\omega^2 kT}[\rho^2(\omega) + \frac{\hbar\omega^3}{\pi^2 c^3}\rho(\omega)]. \end{aligned} \tag{1.55}$$

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

$$\rho(\omega) = \frac{\hbar\omega^3/\pi^2 c^3}{e^{\hbar\omega/kT} - 1} \quad \text{(Planck spectrum)}. \tag{1.56}$$

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 $\omega$, like a material oscillator, has a zero-point energy $\frac{1}{2}\hbar\omega$ (see Chapter 2). The total zero-point energy of a linear dipole oscillator of frequency $\omega$ 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

$$\rho_o(\omega) = (\omega^2/\pi^2 c^3)\frac{1}{2}\hbar\omega = \frac{\hbar\omega^3}{2\pi^2 c^3}. \tag{1.57}$$

If we replace $\rho(\omega)$ in (1.54) by $\rho(\omega) + \rho_o(\omega)$, the left side is unchanged:

$$[\rho(\omega) + \rho_o(\omega)] - \frac{\omega}{3}\frac{d}{d\omega}[\rho(\omega) + \rho_o(\omega)] = \rho(\omega) - \frac{\omega}{3}\frac{d\rho}{d\omega}. \tag{1.58}$$

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 side of (1.54) is changed to

$$\begin{aligned} [\rho(\omega) &+ \rho_o(\omega)][U + \tfrac{1}{2}\hbar\omega] = \rho(\omega)U + \tfrac{1}{2}\hbar\omega\rho(\omega) \\ &\quad + \rho_o(\omega)U + \tfrac{1}{2}\hbar\omega\rho_o(\omega) \\ &= \frac{\pi^2 c^3}{\omega^2}[\rho^2(\omega) + \rho_o(\omega)\rho(\omega) + \frac{\hbar\omega^3}{2\pi^2 c^3}\rho(\omega)] + \frac{1}{2}\hbar\omega\rho_o(\omega) \\ &= \frac{\pi^2 c^3}{\omega^2}[\rho^2(\omega) + \frac{\hbar\omega^3}{\pi^2 c^3}\rho(\omega)] + \frac{1}{2}\hbar\omega\rho_o(\omega), \end{aligned} \tag{1.59}$$

where we have used (1.57) and (1.11) in the form $U = (\pi^2 c^3/\omega^2)\rho(\omega)$.

The term $\frac{1}{2}\hbar\omega\rho_o(\omega)$ in (1.59) 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. We shall see later how this comes about, but for now let us just accept it and drop the term $\frac{1}{2}\hbar\omega\rho_o(\omega)$ in (1.59):

$$[\rho(\omega) + \rho_o(\omega)][U + \tfrac{1}{2}\hbar\omega] \to \frac{\pi^2 c^3}{\omega^2}[\rho^2(\omega) + \frac{\hbar\omega^3}{\pi^2 c^3}\rho(\omega)]. \tag{1.60}$$

From (1.54), (1.58), and (1.60), then, we have

$$\rho(\omega) - \frac{\omega}{3}\frac{d\rho}{d\omega} = \frac{\pi^2 c^3}{3\omega^2 kT}[\rho^2(\omega) + \frac{\hbar\omega^3}{\pi^2 c^3}\rho(\omega)], \tag{1.61}$$

which is exactly the Einstein–Stern equation (1.55). The complete spectrum $\rho(\omega) + \rho_o(\omega)$ is then given by equation (1.40).

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 (1.55). 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 now 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 (1.11) 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 (1.11), for instance, then for consistency we must add the spectral energy density $\rho_o(\omega)$ of the zero-point field to $\rho(\omega)$ :

$$\rho(\omega) + \rho_o(\omega) = \frac{\omega^2}{\pi^2 c^3}(U + \frac{1}{2}\hbar\omega), \tag{1.62}$$

or again $\rho_o(\omega) = \hbar\omega^3/2\pi^2c^3$, 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 (1.11). This is not surprising, for if Planck had simply invoked equipartition of energy and used $U = kT$ in (1.11), 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."[5]

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 $\rho(\omega) + \rho_o(\omega)$ 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 $\rho(\omega) + \rho_o(\omega)$. Crucial to the derivation of the Planck spectrum is the omission of the term $\frac{1}{2}\hbar\omega\rho_o(\omega)$ in (1.59). This omission occurs automatically in the quantum theory of the Einstein–Hopf model, as we shall see in the next chapter. 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 (1916).

[5] *Albert Einstein: Philosopher-Scientist*, ed. P. A. Schilpp (Tudor, New York, 1949), p. 43.



## 1.7 Einstein's Fluctuation Formula

Prior to his work with Hopf and Stern, Einstein (1909) had derived a formula for the energy fluctuations of thermal radiation. 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

$$\langle \Delta E_\omega^2 \rangle = [\hbar\omega\rho(\omega) + \frac{\pi^2 c^3}{\omega^2}\rho^2(\omega)]V\,d\omega. \tag{1.63}$$

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 $h\nu$" :

$$\langle \Delta E_\omega^2 \rangle_{\text{particles}} = \hbar\omega\rho(\omega)V\,d\omega, \tag{1.64}$$

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

$$\langle \Delta E_\omega^2 \rangle_{\text{waves}} = \frac{\pi^2 c^3}{\omega^2}\rho^2(\omega)V\,d\omega. \tag{1.65}$$

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.

The "wave" term (1.65) may be derived from the superposition (1.49) of waves with independent random phases. For instance,

$$\begin{aligned}\langle \mathbf{E}^2(\mathbf{r},t)\rangle \;=\; & -2\mathrm{Re}\sum_{\mathbf{k}_1\lambda_1}\sum_{\mathbf{k}_2\lambda_2}[A_{\mathbf{k}_1\lambda_1}A_{\mathbf{k}_2\lambda_2}e^{-i(\omega_{k_1}+\omega_{k_2})t}e^{i(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{r}} \\ & \times\langle e^{-i(\theta_{\mathbf{k}_1\lambda_1}+\theta_{\mathbf{k}_2\lambda_2})}\rangle - A_{\mathbf{k}_1\lambda_1}A^*_{\mathbf{k}_2\lambda_2}e^{-i(\omega_{k_1}-\omega_{k_2})t} \\ & \times e^{i(\mathbf{k}_1-\mathbf{k}_2)\cdot\mathbf{r}}\langle e^{-i(\theta_{\mathbf{k}_1\lambda_1}-\theta_{\mathbf{k}_2\lambda_2})}\rangle]\mathbf{e}_{\mathbf{k}_1\lambda_1}\cdot\mathbf{e}_{\mathbf{k}_2\lambda_2},\end{aligned} \tag{1.66}$$

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

$$\langle \mathbf{E}^2(\mathbf{r},t)\rangle = 2\sum_{\mathbf{k}\lambda} \mid A_{\mathbf{k}\lambda} \mid^2 \tag{1.67}$$

and similarly

$$\langle \mathbf{E}^4(\mathbf{r},t)\rangle = 8(\sum_{\mathbf{k}\lambda} \mid A_{\mathbf{k}\lambda} \mid^2)^2, \tag{1.68}$$

so that

$$\langle \mathbf{E}^4(\mathbf{r},t)\rangle - \langle \mathbf{E}^2(\mathbf{r},t)\rangle^2 = 4(\sum_{\mathbf{k}\lambda} \mid A_{\mathbf{k}\lambda} \mid^2)^2 = \langle \mathbf{E}^2(\mathbf{r},t)\rangle^2. \tag{1.69}$$

Since the electromagnetic energy density is proportional to $\langle \mathbf{E}^2\rangle$, it follows from (1.69) that the variance in energy associated with frequency $\omega$ is proportional to $\rho^2(\omega)$. We omit the trivial details of the derivation, which leads directly to equation (1.65).

The "particle" term (1.64) 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

$$E = \sum_{\mathbf{k}\lambda} n_{\mathbf{k}\lambda}\hbar\omega_k, \tag{1.70}$$

so that its variance is

$$\langle \Delta E^2\rangle = \sum_{\mathbf{k}\lambda} \langle \Delta n_{\mathbf{k}\lambda}\rangle^2 \hbar^2 {\omega_k}^2, \tag{1.71}$$

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

$$\langle \Delta n_{\mathbf{k}\lambda}^2\rangle = \langle n_{\mathbf{k}\lambda}\rangle \tag{1.72}$$

and

$$\langle \Delta E^2\rangle = \sum_{\mathbf{k}\lambda} \langle n_{\mathbf{k}\lambda}\rangle \hbar^2 {\omega_k}^2. \tag{1.73}$$

Since $\rho(\omega)$ is proportional to the average number of photons at frequency $\omega$, equation (1.73) leads easily to the particle term (1.64) in the Einstein fluctuation formula.

The Einstein fluctuation formula is derived more thoroughly in the next chapter. 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 (1.65) by $\rho(\omega) + \rho_o(\omega)$, where the spectral energy density $\rho_o(\omega)$ of the zero-point field is given by (1.57), we have

$$\begin{aligned}
\langle \Delta E_\omega^2\rangle_{\text{waves}} &\rightarrow \frac{\pi^2 c^3}{\omega^2}[\rho^2(\omega) + 2\rho_o(\omega)\rho(\omega) + \rho_o^2(\omega)]V\,d\omega \\
&= [\frac{\pi^2 c^3}{\omega^2}\rho^2(\omega) + \hbar\omega\rho(\omega)]V\,d\omega + \frac{\pi^2 c^3}{\omega^2}\rho_o^2(\omega)V\,d\omega \\
&= \langle \Delta E_\omega^2\rangle_{\text{waves}} + \langle \Delta E_\omega^2\rangle_{\text{particles}} + \frac{\pi^2 c^3}{\omega^2}\rho_o^2(\omega)V\,d\omega \\
&= \langle \Delta E_\omega^2\rangle_{\text{waves}} + \langle \Delta E_\omega^2\rangle_{\text{particles}} + \frac{1}{2}\hbar\omega\rho_o(\omega)V\,d\omega.
\end{aligned} \tag{1.74}$$

The "extra" (third) term in this expression does not appear in the Einstein fluctuation formula. Indeed we shall see in the following chapter that it does not appear at all in quantum theory, for the same reason that the term $\frac{1}{2}\hbar\omega\rho_o(\omega)$ in equation (1.59) 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 (without zero-point energy). If there were only classical wave fluctuations in thermal radiation, we could ignore the term proportional to $\rho(\omega)$ in equation (1.61). The result is

$$\rho(\omega) - \frac{\omega}{3}\frac{d\rho}{d\omega} = \frac{\pi^2 c^3}{3\omega^2 kT}\rho^2(\omega) \tag{1.75}$$

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, (1.61) becomes

$$\rho(\omega) - \frac{\omega}{3}\frac{d\rho}{d\omega} = \frac{\hbar\omega}{3kT}\rho(\omega) \tag{1.76}$$

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

## 1.8 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 his "heuristic principle" of 1905, and the basis it provided for an "astonishingly simple" derivation of the Planck spectrum.

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 $\rho(\omega_o)$ at the Bohr transition frequency $\omega_o = (E_2 - E_1)/\hbar$ :

$$(\dot{N}_1)_{\text{absorption}} = -B_{12}N_1\rho(\omega_o). \tag{1.77}$$

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 $\omega_o$. One is spontaneous emission, which can occur in the absence of any radiation and is described by the rate constant $A_{21}$ :

$$(\dot{N}_1)_{\text{spontaneous emission}} = A_{21}N_2. \tag{1.78}$$

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

$$(\dot{N}_1)_{\text{stimulated emission}} = B_{21}N_2\rho(\omega_o). \tag{1.79}$$

The condition for equilibrium is

$$(\dot{N}_1)_{\text{absorption}} + (\dot{N}_1)_{\text{spontaneous emission}} + (\dot{N}_1)_{\text{stimulated emission}} = 0 \tag{1.80}$$

or

$$A_{21}N_2 + B_{21}N_2\rho(\omega_o) = B_{12}N_1\rho(\omega_o), \tag{1.81}$$

$$\rho(\omega_o) = \frac{A_{21}/B_{21}}{(B_{12}/B_{21})(N_1/N_2) - 1} = \frac{A_{21}/B_{21}}{(B_{12}/B_{21})e^{\hbar\omega_o/kT} - 1}\,, \tag{1.82}$$

since $N_2/N_1 = e^{-(E_2-E_1)/kT} = e^{-\hbar\omega_o/kT}$ in thermal equilibrium. We are using Bohr's postulate (1913) that $E_2 - E_1 = \hbar\omega_o$ , 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 $\rho(\omega_o)$ becomes so large that spontaneous emission is much less probable than stimulated emission. Then from (1.81) we must have $B_{21} = B_{12}$ and, from (1.82),

$$\rho(\omega_o) = \frac{A_{21}/B_{21}}{e^{\hbar\omega_o/kT} - 1}\,. \tag{1.83}$$

For $kT >> \hbar\omega_o$, furthermore,

$$\rho(\omega_o) \cong \frac{A_{21}}{B_{21}}\frac{kT}{\hbar\omega_o}. \tag{1.84}$$

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

$$\frac{A_{21}}{B_{21}} = \frac{\hbar\omega_o^3}{\pi^2c^3} \tag{1.85}$$

and equation (1.83) then yields the Planck spectrum for $\rho(\omega)$.

This derivation of the Planck spectrum joined 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 has 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, (1.81) and (1.82) are replaced by

$$A_{21}N_2 = B_{12}N_1\rho(\omega_o), \tag{1.86}$$

$$\rho(\omega_o) = \frac{A_{21}}{B_{12}}\frac{N_2}{N_1} = \frac{\hbar\omega_o^3}{\pi^2c^3}e^{-\hbar\omega_o/kT}. \tag{1.87}$$

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." 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 $h\nu$.[6] 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 $\delta t$ it acquires some linear momentum $\Delta$ due to emission and absorption of radiation. Each emission or absorption process imparts to the atom a linear momentum $\lambda_i$, which may be positive or negative. If $n$ emission and absorption processes occur during the time interval $\delta t$, then

$$\Delta = \sum_{i=1}^{n} \lambda_i \tag{1.88}$$

and, assuming the $\lambda_i$ to be independent random variables of zero mean,

$$\langle \Delta^2 \rangle = \sum_{i=1}^{n} \langle \lambda_i^2 \rangle \to \frac{1}{3}\left(\frac{\hbar\omega_o}{c}\right)^2 n \tag{1.89}$$

if we associate with each process of emission or absorption a momentum transfer (photon momentum) $\hbar\omega_o/c$. We have also included a factor of 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 occuring in the time interval $\delta t$ is given, according to the foregoing analysis, by

$$n = N_2 A_{21}\delta t + (N_1 + N_2)B_{12}\rho(\omega_o)\delta t, \tag{1.90}$$

[6] The term *photon* for radiation quanta was coined in 1926 by Gilbert Lewis, a physical chemist.



so that

$$\begin{aligned}(\delta t)^{-1}\langle \Delta^2 \rangle &= \frac{1}{3}\left(\frac{\hbar\omega_o}{c}\right)^2 [N_2 A_{21} + (N_1 + N_2)B_{12}\rho(\omega_o)] \\ &= \frac{2}{3}\left(\frac{\hbar\omega_o}{c}\right)^2 N_1 B_{12}\rho(\omega_o),\end{aligned} \tag{1.91}$$

where we have used the equilibrium condition (1.81).

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

$$F = -Rv = -\left(\frac{\hbar\omega_o}{c^2}\right)(N_1 - N_2)B_{12}\left[\rho(\omega_o) - \frac{\omega_o}{3}\frac{d\rho}{d\omega_o}\right]v. \tag{1.92}$$

As in the classical Einstein–Hopf model the condition for thermal equilibrium is $\langle \Delta^2 \rangle/\delta t = 2RkT$ or, from (1.91) and (1.92),

$$\begin{aligned}\rho(\omega_o) - \frac{\omega_o}{3}\frac{d\rho}{d\omega_o} &= \left(\frac{\hbar\omega_o}{3kT}\right)\left(\frac{N_1}{N_1 - N_2}\right)\rho(\omega_o) \\ &= \left[\frac{\hbar\omega_o/3kT}{1 - e^{-\hbar\omega_o/kT}}\right]\rho(\omega_o).\end{aligned} \tag{1.93}$$

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."

## 1.9 Discussion

In Section 1.6 we alluded to the fact that an oscillator (or atom) in its ground state does not absorb zero-point electromagnetic radiation. The reason for this is discussed in Chapter 4. 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 the preceding section, 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

$$(\dot{N}_2)^{(o)}_{\text{stimulated emission}} = -B_{21}\rho_o(\omega_o)N_2 = -B_{21}\left(\frac{\hbar\omega_o^3}{2\pi^2 c^3}\right)N_2, \tag{1.94}$$

where we have used equation (1.57) for the spectral energy density $\rho_o(\omega)$ of the zero-point field. Using (1.85), therefore, we have

$$(\dot{N}_2)^{(o)}_{\text{stimulated emission}} = -B_{21}\left(\frac{A_{21}}{2B_{21}}\right)N_2 = -\frac{1}{2}A_{21}N_2$$
$$= \frac{1}{2}(\dot{N}_2)_{\text{spontaneous emission}}. \qquad (1.95)$$

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 is induced by the zero-point electromagnetic field." We attempt to clarify the situation in the following chapters.

The result (1.95), 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

$$\frac{N_1}{N_1 - N_2} = 1 + \frac{B_{21}}{A_{21}}\rho(\omega_o), \qquad (1.96)$$

which follows from (1.81), in equation (1.93):

$$\rho(\omega_o) - \frac{\omega_o}{3}\frac{d\rho}{d\omega_o} = \frac{\hbar\omega_o}{3kT}\left[1 + \frac{B_{21}}{A_{21}}\rho(\omega_o)\right]\rho(\omega_o)$$
$$= \frac{\pi^2 c^3}{3\omega_o^2 kT}\left[\rho^2(\omega_o) + \frac{A_{21}}{B_{21}}\rho(\omega_o)\right]. \qquad (1.97)$$

The identity (1.85) shows that this result is equivalent to (1.61). 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 a consequence of spontaneous emission.* The fact that the particle term may also be related as in Section 1.7 to the zero-point field thus suggests again some connection between spontaneous emission and the zero-point field. This connection will be explored in Chapter 4.

We noted in Section 1.7 that the particle term was the nonclassical feature of the Einstein fluctuation formula. In fact this term, which we have just related to the existence of spontaneous emission and zero-point radiation, led Einstein in 1917 to conclude that "Outgoing radiation in the form of spherical waves does not exist."

To understand this conclusion, let us first note that the recoil associated with spontaneous emission contributes only to the particle term in $(\delta t)^{-1}\langle\Delta^2\rangle$, not to the wave term. Now the wave term has contributions



from both absorption and stimulated emission (neither of which contributes to the particle term), and it is obvious from the classical wave picture that absorption and stimulated emission must cause the atom to recoil, simply because the field carries linear momentum. But why does spontaneous emission not contribute likewise to the recoil associated with the wave term?

The reason is simple. In a classical wave description of spontaneous emission, the radiation is a wave with inversion symmetry about the position $\mathbf{r} = 0$ of the atom. Thus any recoil associated with radiation propagating in the direction $\mathbf{r}$ from the atom is cancelled by the contribution from the radiation in the direction $-\mathbf{r}$. The classical wave pattern associated with spontaneous emission is, loosely speaking, "everywhere at once," and its inversion symmetry precludes any possibility of atomic recoil. In the quantum-electrodynamical description of spontaneous emission, however, the radiated field amplitude has the same spatial distribution predicted classically, but it represents a *probability amplitude for directional photon emission.* The expectation value of the net recoil vanishes because there is no preferred direction of emission, just as predicted by the classical wave picture. But contrary to the classical wave picture, there *is* a nonvanishing mean-square momentum transfer to the atom that, for radiation of frequency $\omega_o$, is $(\hbar\omega_o/c)^2$. It is in this sense that the classical picture of outgoing waves fails.

It is perhaps worth noting that the recoil of a spontaneously emitting atom is an experimental fact, as are the recoils associated with the absorption and stimulated emission of radiation. In absorption the recoil is in the same direction as the incoming (absorbed) photon, whereas in stimulated emission the recoil is in the direction opposite to that of the incoming (stimulating) photon; these are simple consequences of the conservation of linear momentum. In spontaneous emission the direction of recoil cannot be predicted, since the direction of the emitted photon is unpredictable. Recoil accompanying spontaneous emission was inferred experimentally by Frisch in 1933, and has in recent years been confirmed more accurately.

We conclude the present discussion with a tribute to the unsung experimentalists who so painstakingly measured blackbody spectra: when Planck fit his formula to their data he obtained $h = 6.55 \times 10^{-27}$ erg-sec for his constant, within 1% of the modern value $h = 6.63 \times 10^{-27}$. For the Boltzmann constant Planck obtained $k = 1.35 \times 10^{-16}$erg/K, the modern value being $1.38 \times 10^{-16}$. (Since the universal gas constant $R = N_A k$ was known, Planck also obtained an accurate estimate of Avogadro's number.)

## 1.10 Specific Heats

It was Maxwell, in 1859, who first suggested that classical physics was wrong. What he later called "the greatest difficulty yet encountered by the molecular theory" had to do with the theory of specific heats of gases.

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

$$c_v = 3N_A k = 3R \approx 6 \text{ cal/mole–K} \quad \text{(Dulong–Petit law)}. \tag{1.98}$$

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 $c_v \to 0$ as $T \to 0$, contradicting the classical prediction (1.98) based on the equipartition theorem.

It was found in 1840 that the specific heat of diamond is smaller than 6 cal/mole-K even at room temperature. This anomaly was first explained by Einstein in 1907. Einstein argued that Planck's equation (1.29) gives the average energy in thermal equilibrium of each (harmonic) vibrational degree of freedom, so that[7]

$$U = \frac{3Nh\nu}{e^{h\nu/kT} - 1} \tag{1.99}$$

and

$$c_v = 3R\left(\frac{\theta}{T}\right)^2 \frac{e^{\theta/T}}{(e^{\theta/T} - 1)^2} \tag{1.100}$$

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 ($T >> \theta$), equation (1.100) reduces to the Dulong–Petit law. At low temperatures, however, $c_v$ is less than the Dulong–Petit value, and in particular $c_v \to 0$ as $T \to 0$. From a fit to experimental data Einstein deduced that $\theta \approx 1300$ K for diamond. A substance with such a large value of $\theta$ will have a small value of $c_v$ even at room temperature.

[7] Einstein presented a derivation of equation (1.29) using in essence the quantum-statistical formula $U = \sum_{n=0}^{\infty} nh\nu e^{-nh\nu/kT} / \sum_{n=0}^{\infty} e^{-nh\nu/kT}$.



In 1913 Einstein and Stern, in the paper discussed in connection with the blackbody problem in Section 1.6, turned their attention to the specific heats of gases. Their work was motivated by the recent report by Eucken that the molar specific heat for $H_2$ at room temperature was about 5 cal/mole-K, but about 3 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 $\nu$ 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 (1.29):

$$U = \frac{1}{2}I(2\pi\nu)^2 = \frac{h\nu}{e^{h\nu/kT} - 1} . \tag{1.101}$$

The rotational contribution to the specific heat is then

$$c_r = N_A \frac{dU}{dT} = N_A \frac{dU}{d\nu}\frac{d\nu}{dT} = N_A\left(4\pi^2 I\nu\right)\frac{d\nu}{dT} = \left(\frac{2R}{k}\right)p\nu\frac{d\nu}{dT} , \tag{1.102}$$

where $p \equiv 2\pi^2 I$. From equation (1.101) it is clear that $\nu$ is a function of $T$; $d\nu/dT$ follows by differentiation of both sides of that equation with respect to $T$,

$$\frac{d\nu}{dT} = \frac{\nu}{T}\left[1 + \frac{kT}{p\nu^2 + h\nu}\right]^{-1} , \tag{1.103}$$

and it follows from (1.102) that

$$c_r = R\frac{2p\nu^2}{kT}\left[1 + \frac{kT}{p\nu^2 + h\nu}\right]^{-1} , \tag{1.104}$$

where $\nu(T)$ is found by solution of (1.101). The rotational specific heat calculated in this way for the example $p = 2.9 \times 10^{-40}$ g cm$^2$ considered by Einstein and Stern is shown in Figure 1.1. 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 equation (1.101) is modified to include zero-point energy:

$$U = p\nu^2 = \frac{h\nu}{e^{h\nu/kT} - 1} + \frac{1}{2}h\nu. \tag{1.105}$$

Following the same steps leading from (1.101) to (1.104), it is found that

$$c_r = R\frac{2p\nu^2}{kT}\left[1 + \frac{kT}{p\nu^2 - h^2/4p}\right]^{-1} , \tag{1.106}$$

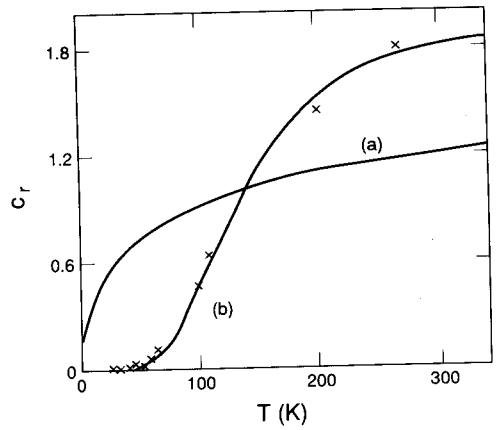


Figure 1.1: Experimental data (×) of Eucken on specific heat of molecular hydrogen; specific heat computed by Einstein and Stern (a) without zero-point energy [equation (1.104)] and (b) with zero-point energy [equation (1.106)].

where $\nu(T)$ is obtained by solving (1.105) for $\nu$ in terms of $T$. The resulting $c_r$ plotted in Figure 1.1 is seen to agree very well with Eucken's observations. At high temperatures $c_r$ asymptotes to $R \approx 2$ cal/mole-K, but at low temperatures $c_r \to 0$.

Einstein and Stern thus gave a very interesting interpretation of Eucken's observation that the specific heat of $H_2$ decreased from 5 cal/mole-K to 3 as $T$ decreased from 300 to 60 K: because of zero-point energy, the rotational contribution to the specific heat decreases from 2 cal/mole-K to 0 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] probable."

The Einstein–Stern explanation turned out to be incorrect. The rotational energy levels of a diatomic molecule are given in quantum theory by $E_J \cong BJ(J+1)$, where $B$ is a constant characteristic of the molecule and $J = 0, 1, 2, \ldots$ Therefore 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 $c_r \to 0$ as $T \to 0$ is due simply to the fact that discrete energy levels are 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 $T \to 0$.

As a consequence of the Einstein–Stern paper, the concept of zero-point energy began to take on greater importance, especially among physical chemists. This was due in part to the growing interest at the time in low-temperature phenomena. Stern himself in 1913 used zero-point energy in a calculation of the vapor pressure of solids.

## 1.11 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. Debye also showed that if Planck's zero-point energy were real, there should be such a reduction in intensity even as $T \to 0$. 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

$$E_s(\mathbf{r}) = \sum_n \frac{p_n}{|\mathbf{r}-\mathbf{r_n}|} e^{-i\omega(t-|\mathbf{r}-\mathbf{r}_n|)/c} = e^{-i\omega t} \sum_n \frac{p_n}{|\mathbf{r}-\mathbf{r_n}|} e^{ik|\mathbf{r}-\mathbf{r}_n|} \quad . \tag{1.108}$$

For distances large compared with the dimensions of the scattering volume we have

$$\begin{aligned} |\mathbf{r}-\mathbf{r}_n| &= [r^2 - 2\mathbf{r}\cdot\mathbf{r}_n + r_n^2]^{1/2} = r[1 - 2\mathbf{r}\cdot\mathbf{r}_n/r^2 + r_n^2/r^2]^{1/2} \\ &\cong r[1 - \mathbf{r}\cdot\mathbf{r}_n/r^2] = r - \mathbf{r}\cdot\mathbf{r}_n/r \ , \end{aligned} \tag{1.109}$$

so that $k|\mathbf{r}-\mathbf{r}_n| \cong kr - (k\mathbf{r}/r)\cdot\mathbf{r}_n \equiv kr - \mathbf{k}\cdot\mathbf{r}_n$ in the exponential in (1.108), where $\mathbf{k}$ is the wave vector of the (elastically) scattered wave. Thus

$$E_s(\mathbf{r}) \cong \frac{1}{r} e^{-i\omega(t-r/c)} \sum_n p_n e^{-i\mathbf{k}\cdot\mathbf{r}_n} \quad . \tag{1.110}$$

We take the strength $p_n$ of the nth scatterer to be proportional to the field $E_o e^{i\mathbf{k}_o\cdot\mathbf{r}_n}$ ($k_o = k$) incident upon it: $p_n = \alpha E_o e^{i\mathbf{k}_o\cdot\mathbf{r}_n}$ and

$$E_s(\mathbf{r}) \cong \frac{\alpha}{r} E_o e^{-i(\omega t - kr)} \sum_n e^{-i\mathbf{K}\cdot\mathbf{r}_n} \quad , \tag{1.111}$$

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

For a periodic lattice of scatterers the scattered field (1.111) 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 $K = [k^2 + k_o^2 - 2\mathbf{k}_o\cdot\mathbf{k}]^{1/2} = [2k^2 - 2k^2\cos 2\theta]^{1/2} = 2k\sin\theta = (4\pi/\lambda)\sin\theta$, the condition that $K$ belongs to the reciprocal lattice is just the Bragg condition, $2d\sin\theta = n\lambda$, where $2\theta$ is the angle between the incident and scattered (diffracted) waves.

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

$$\sum_n e^{-i\mathbf{K}\cdot\mathbf{r}_n} \to e^{-i\mathbf{K}\cdot\mathbf{u}} \sum_n e^{-i\mathbf{K}\cdot\mathbf{r}_n} \, . \tag{1.112}$$

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

$$\langle e^{-i\mathbf{K}\cdot\mathbf{u}}\rangle = 1 - i\mathbf{K}\cdot\langle\mathbf{u}\rangle - \frac{1}{2}\langle(\mathbf{K}\cdot\mathbf{u})^2\rangle + \ldots = 1 - \frac{1}{6}K^2\langle\mathbf{u}^2\rangle + \ldots \tag{1.113}$$

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

$$\langle e^{-i\mathbf{K}\cdot\mathbf{u}}\rangle = e^{-K^2\langle\mathbf{u}^2\rangle/6} \tag{1.114}$$

and $\frac{1}{2}m\omega_o^2\langle\mathbf{u}^2\rangle = \frac{3}{2}kT$, where $m$ and $\omega_o$ 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

$$|\langle e^{-i\mathbf{K}\cdot\mathbf{u}}\rangle|^2 \equiv e^{-2W} = e^{-K^2kT/2m\omega_o^2}. \tag{1.115}$$

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 $T \to 0$ there is a nonvanishing $\langle\mathbf{u}^2\rangle$ associated with zero-point energy:

$$m\omega_o^2\langle\mathbf{u}^2\rangle = 3(\tfrac{1}{2}\hbar\omega_o), \tag{1.116}$$



so that

$$e^{-2W} = e^{-\hbar K^2/2m\omega_o} \quad \text{for } T \to 0. \tag{1.117}$$

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

## 1.12 Molecular Vibrations

Direct evidence for the reality of zero-point energy was provided by Mulliken in 1924. 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 $E_n = \hbar\omega[(n+\frac{1}{2}) - x_e(n+\frac{1}{2})^2 + y_e(n+\frac{1}{2})^3 + \ldots]$, where the constants $\omega, x_e, y_e, \ldots$ are characteristic of the particular molecule, $n = 0, 1, 2, \ldots$, and the zero-point contributions are included. The vibrational frequencies are given by $|E_n - E_{n'}|/\hbar$. Mulliken studied the two molecules $B^{10}O^{16}$ and $B^{11}O^{16}$. 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.

## 1.13 Summary

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 the 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. The great simplicity of Einstein's derivation, perhaps, 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. The role of the zero-point electromagnetic field in spontaneous emission and other electromagnetic processes will be discussed in much greater detail in the following chapters.

We have described how zero-point energy appeared and was used during the development of quantum theory.[8] Although interest in the concept in connection with blackbody 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 1925–26, and long before it was to become central to the world-view of modern physicists.

## 1.14 Bibliography

Debye, P., "Interferenz von Röntgenstrahlen und Wärmebewegung," *Ann. d. Phys.* **43**, 49 (1914).

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

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

Einstein, A., "Zur gegenwärtigen Stand des Strahlungsproblems," *Phys. Zs.* **10**, 185 (1909).

Einstein, A., "Zur Quantentheorie der Strahling," *Phys. Zs.* **18**, 121 (1917).

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

Einstein, A. and L. Hopf, "Statistische Untersuchung der Bewegung eines Resonators in einem Strahlungsfeld," *Ann. d. Phys.* **33**, 1105 (1910b).

Einstein, A. and O. Stern, "Einige Argumente für die Annahme einer molekularen Agitation beim absoluten Nullpunkt," *Ann. d. Phys.* **40**, 551 (1913).

Einstein, A., "Zur Quantentheorie der Strahlung," *Phys. Zs.* **18**, 121 (1917).

Frisch, O. R., "Experimenteller Nachweis des Einsteinschen Strahlungsrückstoßes," *Z. Phys.* **86**, 42 (1933).

Klein, M. J., "Max Planck and the Beginnings of Quantum Theory," *Arch. Hist. Exact Sciences* **1**, 459 (1962).

Klein, M. J., "Thermodynamics and Quanta in Planck's Work," *Physics Today* (November 1966), 23.

Kuhn, T. S., *Black-Body Theory and the Quantum Discontinuity, 1894-1912* (Oxford University Press, New York, 1978).

Mulliken, R. S., "The Band Spectrum of Boron Monoxide," *Nature* **114**, 349 (1924).

Nernst, W., "Über einen Versuch von quantentheoretischen Betrachtungen zur Annahme stetiger Energieänderungen zurückzukehren, " *Verhandl. Deut. Phys. Gen.* **18**, 83 (1916).

Pais, A., *'Subtle is the Lord ...' The Science and the Life of Albert Einstein* (Oxford University Press, New York, 1982).

Planck, M., "Über das Gesetz der Energieverteilung im Normal- spectrum," *Ann. d. Phys.* **4**, 553 (1901).

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

[8] Our survey is not exhaustive. We have not discussed, for instance, the necessity of zero-point energy in Stern's calculation of the vapor pressure of solids.

# Chapter 2

# The Electromagnetic Vacuum

In five minutes you will say that it is all so absurdly simple.
– Sherlock Holmes, "The Adventure of the Dancing Men"
Arthur Conan Doyle

## 2.1 Introduction

The quantum theory of the free electromagnetic field in the absence of any sources was formulated by Born, Heisenberg, and Jordan (1926) in one of the founding papers of quantum theory. The first application was made by Dirac (1927), who treated the emission and absorption of radiation. The new quantum electrodynamics (QED) predicted a fluctuating zero-point or "vacuum" field existing even in the absence of any sources. In this chapter we consider the quantization of the electromagnetic field, with particular emphasis on the vacuum state.

According to contemporary physics the universe is made up of matter fields, whose quanta are fermions (e.g., electrons and quarks), and force fields, whose quanta are bosons (e.g., photons and gluons). *All these fields have zero-point energy.* The oldest and best-known quantized force field is the electromagnetic one. It is important for us to understand the main features of the quantized electromagnetic field, not only because quantum electrodynamics is "the best theory we have," but also because it is in many ways characteristic of all quantum field theories.

## 2.2 The Harmonic Oscillator

A monochromatic electromagnetic field is mathematically equivalent to a harmonic oscillator of the same frequency. Before showing this we will briefly review the harmonic oscillator in quantum mechanics.

The Hamiltonian has the same form as in classical mechanics:

$$H = p^2/2m + \frac{1}{2}m\omega^2 q^2, \tag{2.1}$$

where now $q$ and $p$ are quantum-mechanical operators in a Hilbert space. The Heisenberg equations of motion have the same form as the classical Hamilton equations:

$$\dot{q} = (i\hbar)^{-1}[q, H] = p/m, \tag{2.2}$$

$$\dot{p} = (i\hbar)^{-1}[p, H] = -m\omega^2 q. \tag{2.3}$$

These follow from the commutation rule $[q,p] \equiv qp - pq = i\hbar$. We define the (non-Hermitian) operator

$$a = \frac{1}{\sqrt{2m\hbar\omega}}(p - im\omega q) \tag{2.4}$$

and its adjoint

$$a^\dagger = \frac{1}{\sqrt{2m\hbar\omega}}(p + im\omega q), \tag{2.5}$$

or equivalently

$$q = i\sqrt{\frac{\hbar}{2m\omega}}(a - a^\dagger), \tag{2.6}$$

$$p = \sqrt{\frac{m\hbar\omega}{2}}(a + a^\dagger). \tag{2.7}$$

From $[q,p] = i\hbar$ it follows that

$$[a, a^\dagger] = 1. \tag{2.8}$$

Equations (2.6)–(2.8) allow us to write the Hamiltonian (2.1) in the form

$$H = \frac{1}{2}\hbar\omega(aa^\dagger + a^\dagger a) = \hbar\omega(a^\dagger a + \frac{1}{2}) \,. \tag{2.9}$$

The energy levels of the harmonic oscillator are thus determined by the eigenvalues of the operator $N \equiv a^\dagger a$. We denote the eigenvalues and (normalized) eigenkets of $N$ by $n$ and $|n\rangle$, respectively:

$$N|n\rangle = n|n\rangle. \tag{2.10}$$



Now $\langle n|N|n\rangle = \langle n|a^\dagger a|n\rangle$ is the scalar product of the vector $a|n\rangle$ with itself. It then follows from (2.10) that $n\langle n|n\rangle = n$ is real and positive.

Consider the effect on the vector $a|n\rangle$ of the operator $N$. Obviously $Na|n\rangle = (aN + [N,a])|n\rangle = na|n\rangle + [N,a]|n\rangle$. But (2.8) implies $[N,a] = -a$, and therefore $Na|n\rangle = (n-1)a|n\rangle$. In other words, if $|n\rangle$ is an eigenstate of $N$ with eigenvalue $n$, then $a|n\rangle$ is an eigenstate of $N$ with eigenvalue $n-1$: $a|n\rangle = C|n-1\rangle$. By taking the norm of both sides of this equation we obtain $|C|^2 = n$, and without any loss of generality we can choose the phase such that $C = \sqrt{n}$ . Thus

$$a|n\rangle = \sqrt{n}|n-1\rangle. \tag{2.11}$$

We find similarly that

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle. \tag{2.12}$$

For obvious reasons $a$ and $a^\dagger$ are called *lowering* and *raising* operators.

We have already noted that the eigenvalues $n \geq 0$. But equation (2.11) shows that we can generate eigenstates with lower and lower eigenvalues by successive applications of the lowering operator $a$. Consistency then requires that $a|n\rangle = 0$ for $n < 1$, and (2.11) indicates that this is satisfied only for $n = 0$. The eigenvalues $n$ of $N = a^\dagger a$ are therefore zero and all the positive integers. That is, the energy levels of the harmonic oscillator are given by

$$E_n = (n + \frac{1}{2})\hbar\omega, \quad n = 0, 1, 2, \ldots \tag{2.13}$$

Let us briefly connect this operator approach to that based on the Schrödinger equation in the coordinate representation. From (2.11) we have $a|0\rangle = 0$, or $(p - im\omega q)|0\rangle = 0$ for the ground state $|0\rangle$. Thus $\langle q|(p - im\omega q)|0\rangle = \langle q|p|0\rangle - im\omega q\langle q|0\rangle = 0$. Now $\langle q|0\rangle$ is the wave function $\psi_0(q)$ and $\langle q|p|0\rangle = (\hbar/i)\partial\psi_0/\partial q$, so that

$$(\frac{\hbar}{i}\frac{\partial}{\partial q} - im\omega q)\psi_0(q) = 0 \tag{2.14}$$

or $\psi_0(q) = (m\omega/\pi\hbar)^{1/4}e^{-m\omega q^2/2\hbar}$ when normalized such that $\int_{-\infty}^{\infty} dq|\psi_0(q)|^2 = 1$. The excited-state eigenfunctions $\psi_n(q)$ may be obtained by application of $a^\dagger$ according to (2.12): $|n\rangle = (n!)^{-1/2}(a^\dagger)^n|0\rangle$ and

$$\begin{aligned}
\psi_n(q) &= \langle q|n\rangle = (n!)^{-1/2}(2m\hbar\omega)^{-n/2}\langle q|(p + im\omega q)^n|0\rangle \\
&= [(2m\hbar\omega)^n n!]^{-1/2}(\frac{\hbar}{i}\frac{\partial}{\partial q} + im\omega q)^n\psi_0(q) \\
&= i^n(2^n n!)^{-1/2}(\frac{m\omega}{\pi\hbar})^{1/4}(\xi - \frac{\partial}{\partial\xi})^n e^{-\xi^2/2},
\end{aligned} \tag{2.15}$$

where $\xi \equiv (m\omega/\hbar)^{1/2}q$. These eigenstates are proportional to $e^{-\xi^2/2}H_n(\xi)$, where $H_n$ is a Hermite polynomial of degree $n$.

Various properties of the harmonic oscillator can be worked out using either the raising and lowering operators $a$ and $a^\dagger$ or the eigenfunctions $\psi_n(q)$. For instance, we find that $\langle n|q|n\rangle = \langle n|p|n\rangle = 0$ and

$$\langle n|q^2|n\rangle = -\frac{\hbar}{2m\omega}\langle n|(a-a^\dagger)^2|n\rangle = \frac{\hbar}{m\omega}(n+\frac{1}{2}), \tag{2.16}$$

$$\langle n|p^2|n\rangle = m\hbar\omega(n+\frac{1}{2}), \tag{2.17}$$

since $\langle n|m\rangle = \delta_{nm}$. (Recall that eigenkets corresponding to different eigenvalues are orthogonal in the case of a Hermitian operator like $a^\dagger a$.) Thus $\Delta q_n \Delta p_n = (n+\frac{1}{2})\hbar$, where $(\Delta q_n)^2 \equiv \langle n|q^2|n\rangle - \langle n|q|n\rangle^2$. This is consistent with the general uncertainty relation $\Delta q \Delta p \geq \hbar/2$ and shows that the ground state of the harmonic oscillator is a state of minimal uncertainty product. In other words, the ground state is a *coherent state* of the harmonic oscillator.

## 2.3 A Field Mode Is a Harmonic Oscillator

We will now take the most elementary route to the quantization of the electromagnetic field. The first step is to show that a field mode is equivalent to a harmonic oscillator.

The Maxwell equations for the "free" field, i.e., the field in a region where there are no sources, are

$$\nabla \cdot \mathbf{E} = 0, \tag{2.18}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{2.19}$$

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \tag{2.20}$$

$$\nabla \times \mathbf{B} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}. \tag{2.21}$$

We introduce the vector potential $\mathbf{A}$ by writing $\mathbf{B} = \nabla \times \mathbf{A}$. Since $\nabla \cdot (\nabla \times \mathbf{A}) = 0$, (2.19) is automatically satisfied. Equation (2.20) implies $\mathbf{E} = -(1/c)\partial\mathbf{A}/\partial t - \nabla\phi$, where $\phi$ is the scalar potential. From (2.21) we have

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = 0 \tag{2.22}$$



in the Coulomb gauge defined by $\nabla \cdot \mathbf{A} = 0$ and, in the absence of any sources, $\phi = 0$. Obviously (2.18) is then also satisfied. Thus we can obtain a solution of the free-space Maxwell equations by solving (2.22) for the Coulomb-gauge vector potential subject to appropriate boundary conditions.

Separation of variables gives monochromatic solutions

$$\begin{aligned}\mathbf{A}(\mathbf{r},t) &= \alpha(t)\mathbf{A}_o(\mathbf{r}) + \alpha^*(t)\mathbf{A}_o^*(\mathbf{r}) \\ &= \alpha(0)e^{-i\omega t}\mathbf{A}_o(\mathbf{r}) + \alpha^*(0)e^{i\omega t}\mathbf{A}_o^*(\mathbf{r}),\end{aligned} \tag{2.23}$$

where $\mathbf{A}_o(\mathbf{r})$ satisfies the Helmholtz equation,

$$\nabla^2\mathbf{A}_o(\mathbf{r}) + k^2\mathbf{A}_o(\mathbf{r}) = 0 \quad (k = \omega/c), \tag{2.24}$$

and $\alpha(t)$ satisfies $\ddot{\alpha}(t) = -\omega^2\alpha(t)$. The electric and magnetic field vectors are given by

$$\mathbf{E}(\mathbf{r},t) = -\frac{1}{c}[\dot{\alpha}(t)\mathbf{A}_o(\mathbf{r}) + \dot{\alpha}^*(t)\mathbf{A}_o^*(\mathbf{r})], \tag{2.25}$$

$$\mathbf{B}(\mathbf{r},t) = \alpha(t)\nabla \times \mathbf{A}_o(\mathbf{r}) + \alpha^*(t)\nabla \times \mathbf{A}_o^*(\mathbf{r}), \tag{2.26}$$

and the electromagnetic energy is proportional to

$$\begin{aligned}\int d^3r(\mathbf{E}^2+\mathbf{B}^2) &= \frac{1}{c^2}\dot{\alpha}(t)^2\int d^3r\mathbf{A}_o(\mathbf{r})^2 \\ &+ \frac{1}{c^2}\dot{\alpha}^*(t)^2\int d^3r\mathbf{A}_o^*(\mathbf{r})^2 + \frac{2}{c^2}|\dot{\alpha}(t)|^2\int d^3r|\mathbf{A}_o(\mathbf{r})|^2 \\ &+ \alpha(t)^2\int d^3r[\nabla \times \mathbf{A}_o(\mathbf{r})]^2 + \alpha^*(t)^2\int d^3r[\nabla \times \mathbf{A}_o^*(\mathbf{r})]^2 \\ &+ 2|\alpha(t)|^2\int d^3r|\nabla \times \mathbf{A}_o(\mathbf{r})|^2.\end{aligned} \tag{2.27}$$

We show in Appendix C that we may take

$$\int d^3r[\nabla \times \mathbf{A}_o(\mathbf{r})]^2 = k^2\int d^3r\mathbf{A}_o(\mathbf{r})^2, \tag{2.28}$$

with similar expressions for the terms involving $[\nabla \times \mathbf{A}_o^*(\mathbf{r})]^2$ and $|\nabla \times \mathbf{A}_o(\mathbf{r})|^2$ in (2.27) We also note that $\dot{\alpha}(t)^2 = -\omega^2\alpha(t)^2$, since $\dot{\alpha}(t) = -i\omega\alpha(t)$. Then (2.27) simplifies to

$$H_F = \frac{1}{8\pi}\int d^3r(\mathbf{E}^2+\mathbf{B}^2) = \frac{k^2}{2\pi}|\alpha(t)|^2, \tag{2.29}$$

where, without any loss of generality, we assume the "mode function" $\mathbf{A}_o(\mathbf{r})$ is normalized such that

$$\int d^3r|\mathbf{A}_o(\mathbf{r})|^2 = 1. \tag{2.30}$$

Define the real quantities

$$q(t) = \frac{i}{c\sqrt{4\pi}}[\alpha(t) - \alpha^*(t)], \tag{2.31}$$

$$p(t) = \frac{k}{\sqrt{4\pi}}[\alpha(t) + \alpha^*(t)], \tag{2.32}$$

in terms of which equation (2.29) is

$$H_F = \frac{1}{2}(p^2 + \omega^2 q^2). \tag{2.33}$$

The notation suggests that our field mode of frequency $\omega$ is mathematically equivalent to a harmonic oscillator of frequency $\omega$. To prove this we must, of course, show that $q$ and $p$ are indeed canonically conjugate coordinate and momentum variables. But this is trivial: from the definitions (2.31) and (2.32) and $\dot{\alpha} = -i\omega\alpha$, we have $\dot{q} = p$ and $\dot{p} = -\omega^2 q$, which are the Hamilton equations that follow from the Hamiltonian $H_F$.

## 2.4 Quantization of a Field Mode

To describe a field mode quantum mechanically, we simply describe the equivalent harmonic oscillator quantum mechanically. Since the oscillator with Hamiltonian (2.33) has unit mass, we introduce raising and lowering operators $a$ and $a^\dagger$ using (2.6) and (2.7) with $m = 1$. Comparing with (2.31) and (2.32), we see that this quantization procedure is equivalent to replacing the classical variable $\alpha(t)/c\sqrt{4\pi}$ by the quantum-mechanical operator $(\hbar/2\omega)^{1/2}a(t)$, or $\alpha(t)$ by $(2\pi\hbar c^2/\omega)^{1/2}a(t)$ and $\alpha^*(t)$ by $(2\pi\hbar c^2/\omega)^{1/2}a^\dagger(t)$. That is, except for trivial constants that depend upon the arbitrary normalization chosen for the mode function, $\alpha(t)$ and $\alpha^*(t)$ in the classical theory are replaced by the lowering and raising operators $a(t)$ and $a^\dagger(t)$, respectively, in quantum theory.

The classical vector potential (2.23) is thus replaced by the *operator*

$$\mathbf{A}(\mathbf{r},t) = \left(\frac{2\pi\hbar c^2}{\omega}\right)^{1/2}[a(t)\mathbf{A}_o(\mathbf{r}) + a^\dagger(t)\mathbf{A}_o^*(\mathbf{r})], \tag{2.34}$$



and the operators corresponding to the electric and magnetic fields are similarly

$$\mathbf{E}(\mathbf{r},t) = i(2\pi\hbar\omega)^{1/2}[a(t)\mathbf{A}_o(\mathbf{r}) - a^\dagger(t)\mathbf{A}_o^*(\mathbf{r})], \tag{2.35}$$

$$\mathbf{B}(\mathbf{r},t) = \left(\frac{2\pi\hbar c^2}{\omega}\right)^{1/2}[a(t)\nabla\times\mathbf{A}_o(\mathbf{r}) + a^\dagger(t)\nabla\times\mathbf{A}_o^*(\mathbf{r})]. \tag{2.36}$$

The Hamiltonian (2.33) for the quantized field mode is now obviously equivalent to

$$H_F = \hbar\omega(a^\dagger a + \frac{1}{2}). \tag{2.37}$$

The energy eigenvalues of a field mode of frequency $\omega$ are given by equation (2.13). The integer $n$ is the number of energy quanta or *photons* in the field mode described by the state $|n\rangle$. The vacuum state $|0\rangle$ has no photons, but it nevertheless has an energy $\frac{1}{2}\hbar\omega$. The quantum theory of radiation thus predicts the existence of a zero-point electromagnetic field. In the vacuum state, and in all stationary states $|n\rangle$, the expectation values of the electric and magnetic fields vanish:

$$\langle\mathbf{E}(\mathbf{r},t)\rangle = \langle\mathbf{B}(\mathbf{r},t)\rangle = 0, \tag{2.38}$$

since $\langle n|a|n\rangle = 0$. This means that the electric and magnetic field vectors fluctuate with zero mean in the state $|n\rangle$, although the field has a definite, nonfluctuating energy $(n + \frac{1}{2})\hbar\omega$.

Consider the expectation value of the square of the electric field. From (2.35) this is given by

$$\begin{aligned}\langle\mathbf{E}^2(\mathbf{r},t)\rangle = &-(2\pi\hbar\omega)[\langle a^2(t)\rangle\mathbf{A}_o^2(\mathbf{r}) - \langle a(t)a^\dagger(t) + a^\dagger(t)a(t)\rangle \\ &\times|\mathbf{A}_o(\mathbf{r})|^2 + \langle a^\dagger(t)^2\rangle\mathbf{A}_o^*(\mathbf{r})^2].\end{aligned} \tag{2.39}$$

In the state $|n\rangle$ we have $\langle a^2\rangle = \langle a^{\dagger 2}\rangle = 0$, $\langle aa^\dagger + a^\dagger a\rangle = \langle 2a^\dagger a + 1\rangle = 2n+1$, and

$$\begin{aligned}\langle\mathbf{E}^2(\mathbf{r},t)\rangle &= (n+\frac{1}{2})4\pi\hbar\omega|\mathbf{A}_o(\mathbf{r})|^2 \\ &= 4\pi\hbar\omega|\mathbf{A}_o(\mathbf{r})|^2 n + 2\pi\hbar\omega|\mathbf{A}_o(\mathbf{r})|^2 \\ &\equiv 4\pi\hbar\omega|\mathbf{A}_o(\mathbf{r})|^2 n + \langle\mathbf{E}^2(\mathbf{r})\rangle_0.\end{aligned} \tag{2.40}$$

From the first term on the right we can begin to understand how the quantum theory of radiation resolves the "paradox" of the wave-particle duality of light, for this term, which is a measure of the "intensity" (energy density) of the field at $\mathbf{r}$, has both wave and particle factors. The factor $n$ is

the number of photons in the field, whereas the factor $|\mathbf{A}_o(\mathbf{r})|^2$ gives the same spatial dependence for the intensity predicted by the classical wave theory. Even in the case of a single photon ($n = 1$), the classical wave theory gives the same spatial dependence as quantum theory for the intensity, but this pattern represents the relative *probability distribution* for finding the photon. The detection of a single photon does not produce the spread-out classical intensity pattern $|\mathbf{A}_o(\mathbf{r})|^2$. Instead there is relatively high probability of detecting the photon at points where $|\mathbf{A}_o(\mathbf{r})|^2$ is large, and low probability where $|\mathbf{A}_o(\mathbf{r})|^2$ is small. If $|\mathbf{A}_o(\mathbf{r})|^2 = 0$, the probability of detecting the photon at $\mathbf{r}$ is zero. The wave and particle aspects of the field are thus reconciled by this association of a particle (photon) with the classical (wave) intensity pattern. Comparing (2.25) and (2.26) to (2.35) and (2.36), we can say that the spatial pattern of the field is exactly the same as predicted classically: the quantum mechanics of the field is entirely contained, as it were, in its time dependence.

Quantities like $\langle \mathbf{E}^m(\mathbf{r},t)\rangle$ with $m > 2$ are also easy to calculate. Suppose, for simplicity, that $\mathbf{A}_o^*(\mathbf{r}) = \mathbf{A}_o(\mathbf{r})$, so that

$$\mathbf{E}(\mathbf{r},t) = i(2\pi\hbar\omega)^{1/2}[a(t) - a^\dagger(t)]\mathbf{A}_o(\mathbf{r}) = (4\pi\omega^2)^{1/2}q(t)\mathbf{A}_o(\mathbf{r}). \quad (2.41)$$

Then from the probability distribution $|\psi_0(q)|^2 = (\omega/\pi\hbar)^{1/2}e^{-\omega q^2/\hbar}$ for a ground-state harmonic oscillator of unit mass, we easily obtain the probability distribution

$$P[\mathbf{E}(\mathbf{r},t)] = [2\pi\langle \mathbf{E}^2(\mathbf{r})\rangle_0]^{1/2}\exp\left[-\mathbf{E}^2(\mathbf{r},t)/2\langle \mathbf{E}^2(\mathbf{r})\rangle_0\right] \quad (2.42)$$

for the electric field in the vacuum state $|0\rangle$. Thus $\langle \mathbf{E}^m(\mathbf{r},t)\rangle_0 = 0$ for odd m and

$$\begin{aligned}\langle \mathbf{E}^m(\mathbf{r},t)\rangle_0 &= [2\pi\langle \mathbf{E}^2(\mathbf{r})\rangle_0]^{1/2}\int_0^\infty dE E^m \exp\left[-E^2/2\langle \mathbf{E}^2(\mathbf{r})\rangle_0\right] \\ &= 2^{m/2}\pi^{-1/2}\Gamma\left(\frac{m+1}{2}\right)\langle \mathbf{E}^2(\mathbf{r})\rangle_0^{m/2} \quad \text{(m even)}. \quad (2.43)\end{aligned}$$

Similar results, with the appearance of Hermite polynomials $H_n$, are found for the expectation values of field powers in photon states $|n\rangle$.

What is the physical significance of these vacuum-state expectation values, and in particular of $\langle \mathbf{E}^2(\mathbf{r})\rangle_0$? One thing they indicate is that the electromagnetic vacuum is a stationary *state* of the field *with statistical fluctuations* of the electric and magnetic fields. As far as measurements are concerned, however, it is often argued that the entire universe is evidently bathed in a zero-point electromagnetic field, which can add only



some constant amount to expectation values, as in equation (2.40). Physical measurements will therefore reveal only *deviations* from the vacuum state. Thus the field Hamiltonian (2.37), for example, can be replaced by

$$\begin{aligned}H_{\mathrm{F}} - \langle 0|H_{\mathrm{F}}|0\rangle &= \frac{1}{2}\hbar\omega(aa^\dagger + a^\dagger a) - \frac{1}{2}\hbar\omega \\ &= \frac{1}{2}\hbar\omega(2a^\dagger a + 1) - \frac{1}{2}\hbar\omega \\ &= \hbar\omega a^\dagger a \quad (2.44)\end{aligned}$$

without affecting any physical predictions of the theory. The new Hamiltonian (2.44) is said to be *normally ordered* (or Wick ordered), the raising operator $a^\dagger$ appearing to the left of the lowering operator $a$. The normally ordered Hamiltonian is denoted $:H_{\mathrm{F}}:$, i.e.,

$$:H_{\mathrm{F}}: = :\frac{1}{2}\hbar\omega(aa^\dagger + a^\dagger a): = \hbar\omega a^\dagger a. \quad (2.45)$$

In other words, within the normal ordering symbol we can commute $a$ and $a^\dagger$. Since zero-point energy is intimately connected to the noncommutativity of $a$ and $a^\dagger$, the normal ordering procedure eliminates any contribution from the zero-point field. This is especially reasonable in the case of the field Hamiltonian, since the zero-point term merely adds a constant energy which can be eliminated by a simple redefinition of the zero of energy. Moreover, this constant energy in the Hamiltonian obviously commutes with $a$ and $a^\dagger$ and so cannot have any effect on the quantum dynamics described by the Heisenberg equations of motion.

So the argument goes. However, things are not quite that simple, for in general relativity the zero of energy is not arbitrary. Furthermore we shall see that it is possible to attribute measurable effects, such as the Casimir force and the Lamb shift, to *changes* in zero-point energy. And finally, as discussed in Section 2.6, the zero-point field is not eliminated by dropping its energy from the Hamiltonian.

## 2.5 The Field in Free Space

The generalization of the quantization procedure to a multimode field is straightforward. In this section we consider the field in free space with no physical boundaries, in which case the number of allowed modes is infinite.

Obviously the field intensity for infinite free space should be independent of position so that, from (2.40), $|\mathbf{A}_o(\mathbf{r})|^2$ should be independent of $\mathbf{r}$ for

each mode of the field. Of course $\mathbf{A}_o(\mathbf{r})$ must still satisfy the Helmholtz equation (2.24). A mode function satisfying these conditions is obviously $\mathbf{A}_o(\mathbf{r}) = \mathbf{e}_\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{r}}$, where $\mathbf{k}\cdot\mathbf{e}_\mathbf{k} = 0$ in order to have the transversality condition $\nabla\cdot\mathbf{A}(\mathbf{r},t) = 0$ satisfied for the Coulomb gauge in which we are working.

We also wish to normalize our mode functions according to equation (2.30). To achieve the desired normalization we pretend that space is divided into cubes of volume $V = L^3$ and impose on the field the periodic boundary condition

$$\mathbf{A}(x+L, y+L, z+L, t) = \mathbf{A}(x,y,z,t), \tag{2.46}$$

or equivalently

$$(k_x, k_y, k_z) = \frac{2\pi}{L}(n_x, n_y, n_z), \tag{2.47}$$

where each $n$ can assume any integer value. Of course this artificial periodic boundary condition will be of no physical consequences if $L$ is very large compared with any physical dimensions of interest. It allows us to consider the field in any *one* of the imaginary cubes, and to define a mode function $\mathbf{A}_\mathbf{k}(\mathbf{r}) = V^{-1/2}\mathbf{e}_\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{r}}$ satisfying the Helmholtz equation, transversality, and the "box normalization"

$$\int_V d^3r |\mathbf{A}_\mathbf{k}(\mathbf{r})|^2 = 1, \tag{2.48}$$

where $\mathbf{e}_\mathbf{k}$ is chosen to be a unit vector.

The unit vector $\mathbf{e}_\mathbf{k}$, which we take to be real, specifies the polarization of the field mode. The condition $\mathbf{k}\cdot\mathbf{e}_\mathbf{k} = 0$ means there are two independent choices for $\mathbf{e}_\mathbf{k}$, which we call $\mathbf{e}_{\mathbf{k}1}$ and $\mathbf{e}_{\mathbf{k}2}$, $\mathbf{e}_{\mathbf{k}1}\cdot\mathbf{e}_{\mathbf{k}2} = 0$ and $\mathbf{e}_{\mathbf{k}1}^2 = \mathbf{e}_{\mathbf{k}2}^2 = 1$. Thus we define the mode functions

$$\mathbf{A}_{\mathbf{k}\lambda}(\mathbf{r}) = V^{-1/2}\mathbf{e}_{\mathbf{k}\lambda} e^{i\mathbf{k}\cdot\mathbf{r}} \quad (\lambda = 1,2), \tag{2.49}$$

in terms of which the vector potential (2.34) becomes

$$\mathbf{A}_{\mathbf{k}\lambda}(\mathbf{r},t) = \left(\frac{2\pi\hbar c^2}{\omega_k V}\right)^{1/2} [a_{\mathbf{k}\lambda}(t) e^{i\mathbf{k}\cdot\mathbf{r}} + a_{\mathbf{k}\lambda}^\dagger(t) e^{-i\mathbf{k}\cdot\mathbf{r}}]\mathbf{e}_{\mathbf{k}\lambda}, \tag{2.50}$$

or

$$\mathbf{A}_{\mathbf{k}\lambda}(\mathbf{r},t) = \left(\frac{2\pi\hbar c^2}{\omega_k V}\right)^{1/2} [a_{\mathbf{k}\lambda}(0) e^{-i(\omega_k t - \mathbf{k}\cdot\mathbf{r})} + a_{\mathbf{k}\lambda}^\dagger(0) e^{i(\omega_k t - \mathbf{k}\cdot\mathbf{r})}]\mathbf{e}_{\mathbf{k}\lambda}, \tag{2.51}$$



where $\omega_k = kc$ and $a_{\mathbf{k}\lambda}$, $a_{\mathbf{k}\lambda}^\dagger$ are respectively the *photon annihilation and creation operators* for the mode with wave vector $\mathbf{k}$ and polarization $\lambda$. This gives the vector potential for a plane-wave mode of the field. The condition (2.47) shows that there is an infinite number of such modes. The linearity of Maxwell's equations allows us to write

$$\mathbf{A}(\mathbf{r},t) = \sum_{\mathbf{k}\lambda} \left(\frac{2\pi\hbar c^2}{\omega_k V}\right)^{1/2} [a_{\mathbf{k}\lambda}(t) e^{i\mathbf{k}\cdot\mathbf{r}} + a_{\mathbf{k}\lambda}^\dagger(t) e^{-i\mathbf{k}\cdot\mathbf{r}}]\mathbf{e}_{\mathbf{k}\lambda} \tag{2.52}$$

for the total vector potential in free space.

Using the fact that

$$\int_V d^3r \mathbf{A}_{\mathbf{k}\lambda}(\mathbf{r})\cdot\mathbf{A}_{\mathbf{k}'\lambda'}^*(\mathbf{r}) = \delta_{\mathbf{k},\mathbf{k}'}^3\delta_{\lambda\lambda'}\,, \tag{2.53}$$

we find from the same sort of analysis as in the preceding section that the field Hamiltonian is

$$H_F = \sum_{\mathbf{k}\lambda} \hbar\omega_k (a_{\mathbf{k}\lambda}^\dagger a_{\mathbf{k}\lambda} + \frac{1}{2}) \tag{2.54}$$

for the infinity of modes in free space. This is the Hamiltonian for an infinite number of *uncoupled* harmonic oscillators. Thus the different modes of the field are independent and satisfy the commutation relations

$$[a_{\mathbf{k}\lambda}(t), a_{\mathbf{k}'\lambda'}^\dagger(t)] = \delta_{\mathbf{k},\mathbf{k}'}^3\delta_{\lambda\lambda'} \tag{2.55}$$

and $[a_{\mathbf{k}\lambda}(t), a_{\mathbf{k}'\lambda'}(t)] = [a_{\mathbf{k}\lambda}^\dagger(t), a_{\mathbf{k}'\lambda'}^\dagger(t)] = 0$. From (2.52) it follows that

$$\mathbf{E}(\mathbf{r},t) = i\sum_{\mathbf{k}\lambda} \left(\frac{2\pi\hbar\omega_k}{V}\right)^{1/2} [a_{\mathbf{k}\lambda}(t) e^{i\mathbf{k}\cdot\mathbf{r}} - a_{\mathbf{k}\lambda}^\dagger(t) e^{-i\mathbf{k}\cdot\mathbf{r}}]\mathbf{e}_{\mathbf{k}\lambda}, \tag{2.56}$$

$$\mathbf{B}(\mathbf{r},t) = i\sum_{\mathbf{k}\lambda} \left(\frac{2\pi\hbar c^2}{\omega_k V}\right)^{1/2} [a_{\mathbf{k}\lambda}(t) e^{i\mathbf{k}\cdot\mathbf{r}} - a_{\mathbf{k}\lambda}^\dagger(t) e^{-i\mathbf{k}\cdot\mathbf{r}}]\mathbf{k}\times\mathbf{e}_{\mathbf{k}\lambda}. \tag{2.57}$$

It is worth noting that the free-space mode functions (2.49) form a complete set for transverse vector fields satisfying our periodic boundary condition. That is, the plane-wave modes $\mathbf{A}_{\mathbf{k}\lambda}(\mathbf{r})$ form a complete set in terms of which any mode of the field may be expanded. This is essentially just a statement of Fourier's theorem about the completeness of sines and

cosines. Of course the $\mathbf{A}_{\mathbf{k}\lambda}(\mathbf{r})$ are complete only for modes satisfying the periodic boundary condition, but in a slightly more sophisticated approach we can work with a complete *continuum* of plane-wave mode functions in which the $\mathbf{k}$ vectors are not restricted to the discrete spectrum (2.47) (Chapter 10). This has formal consequences such as the replacement of $\delta^3_{\mathbf{k},\mathbf{k}'}$ in (2.53) and (2.55) by $\delta^3(\mathbf{k}-\mathbf{k}')$, but since it has no physical consequences here, we will just stick to the periodic boundary condition.

The linear momentum of the field is given classically by $\mathbf{P} = (1/4\pi c) \times \int_V d^3r(\mathbf{E}\times\mathbf{B})$. In the case of the quantized field we use (2.56) and (2.57) in this expression and obtain, after straightforward manipulations,

$$\mathbf{P} = \sum_{\mathbf{k}\lambda} \hbar\mathbf{k}(a^\dagger_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} + \frac{1}{2}). \tag{2.58}$$

Obviously $[\mathbf{P}, H_F] = 0$, so that the linear momentum of the field in the absence of any sources is a constant of the motion. It is also obvious that the eigenvalues of $\mathbf{P}$ are $\sum_{\mathbf{k}\lambda} \hbar\mathbf{k}(n_{\mathbf{k}\lambda}+\frac{1}{2})$, where each $n$ is a positive integer or zero. A stationary state of the free field is thus characterized by the set of photon numbers $\{n_{\mathbf{k}\lambda}\}$. The state $|\{n_{\mathbf{k}\lambda}\}\rangle$ has a total photon number $\sum_{\mathbf{k}\lambda} n_{\mathbf{k}\lambda}$, an energy

$$E = \sum_{\mathbf{k}\lambda} \hbar\omega_k(n_{\mathbf{k}\lambda} + \frac{1}{2}), \tag{2.59}$$

and a linear momentum

$$\mathbf{P} = \sum_{\mathbf{k}\lambda} \hbar\mathbf{k}(n_{\mathbf{k}\lambda} + \frac{1}{2}) \tag{2.60}$$

or

$$E = \sum_{\mathbf{k}\lambda} \hbar\omega_k n_{\mathbf{k}\lambda}\ , \tag{2.61}$$

$$\mathbf{P} = \sum_{\mathbf{k}\lambda} \hbar\mathbf{k} n_{\mathbf{k}\lambda} \tag{2.62}$$

if the zero-point energy and linear momentum associated with the vacuum state are discarded. Note that the zero-point momentum $\sum_{\mathbf{k}\lambda} \frac{1}{2}\hbar\mathbf{k}$ in fact vanishes since for each $\mathbf{k}$ there is an equal contribution from $-\mathbf{k}$ in the summation.

We have thus arrived at the quantum theory of the free electromagnetic field in which stationary states are described by photons of energy $\hbar\omega_k$ and linear momentum $\hbar\mathbf{k}$. Since $E^2 - P^2c^2 = \hbar^2(\omega_k^2 - k^2c^2) = 0$ for each photon,



the photons have zero rest mass. The theory also implies that photons are bosons, i.e., that the stationary states are symmetric with respect to permutations of identical photons. To see this, note from equation (2.12) that the $n$-photon state $|n\rangle$ of a field mode may be written in the form

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle, \tag{2.63}$$

which is obviously symmetric with respect to any permutations of the $n$ photons. Of course the boson character of photons is just a consequence of the commutation rule (2.55), from which (2.63) follows.

The $\mathbf{k}$ vector of a photon of mode $(\mathbf{k},\lambda)$ specifies the energy and linear momentum of the photon. The polarization index $\lambda$ is connected with the intrinsic angular momentum, or spin, of the photon. To establish this connection we first note that the intrinsic angular momentum may be defined by the formula[1]

$$\mathbf{M}_s = \frac{1}{4\pi c}\int_V d^3r(\mathbf{E}\times\mathbf{A}). \tag{2.64}$$

From (2.52) and (2.56) we obtain for $\mathbf{M}_s$ the expression

$$\mathbf{M}_s = i\hbar \sum_{\mathbf{k}\lambda} \hat{\mathbf{k}}(a^\dagger_{\mathbf{k}2} a_{\mathbf{k}1} - a^\dagger_{\mathbf{k}1} a_{\mathbf{k}2}), \tag{2.65}$$

where the unit vector $\hat{\mathbf{k}} \equiv \mathbf{k}/k = \mathbf{e}_{\mathbf{k}1} \times \mathbf{e}_{\mathbf{k}2}$. This operator does not commute with $a^\dagger_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda}$, and therefore a photon number state $|n_{\mathbf{k}\lambda}\rangle$ is not an eigenstate of $\mathbf{M}_s$. To construct simultaneous eigenstates of energy, linear momentum, and intrinsic angular momentum of a photon we define the complex unit polarization vectors

$$\mathbf{e}_{\mathbf{k},+1} = -\sqrt{\frac{1}{2}}(\mathbf{e}_{\mathbf{k}1} + i\mathbf{e}_{\mathbf{k}2}), \tag{2.66}$$

$$\mathbf{e}_{\mathbf{k},-1} = \sqrt{\frac{1}{2}}(\mathbf{e}_{\mathbf{k}1} - i\mathbf{e}_{\mathbf{k}2}), \tag{2.67}$$

satisfying $\mathbf{e}^*_{\mathbf{k}\alpha} \cdot \mathbf{e}_{\mathbf{k}\alpha'} = \delta_{\alpha\alpha'}$, $\mathbf{e}^*_{\mathbf{k},\alpha} \times \mathbf{e}_{\mathbf{k},\alpha'} = i\alpha\hat{\mathbf{k}}\delta_{\alpha\alpha'}$, $\alpha = \pm 1$. It is easily seen that, whereas our original polarization vectors $\mathbf{e}_{\mathbf{k}\lambda}$ with $\lambda = 1, 2$

[1] See, for instance, Heitler (1966), Appendix, Section 1. It is worth noting that (2.64) is gauge-invariant, since the vector potential is transverse in the Coulomb gauge employed here, and the transverse part of the vector potential is unaffected by gauge transformations.

correspond to two orthogonal linear polarizations, the new polarization vectors $\mathbf{e}_{\mathbf{k}\alpha}$ with $\alpha = \pm 1$ correspond to opposite *circular* polarizations. We define the photon annihilation operators for the circularly polarized modes $(\mathbf{k}, \alpha)$ by

$$a_{\mathbf{k},+1} = -\sqrt{\frac{1}{2}}(a_{\mathbf{k}1} - ia_{\mathbf{k}2}), \tag{2.68}$$

$$a_{\mathbf{k},-1} = \sqrt{\frac{1}{2}}(a_{\mathbf{k}1} + ia_{\mathbf{k}2}), \tag{2.69}$$

in terms of which

$$\mathbf{M}_s = \hbar \sum_{\mathbf{k}} \hat{\mathbf{k}}(a^\dagger_{\mathbf{k},+1}a_{\mathbf{k},+1} - a^\dagger_{\mathbf{k},-1}a_{\mathbf{k},-1}) = \sum_{\mathbf{k}\alpha} \alpha\hbar\hat{\mathbf{k}}a^\dagger_{\mathbf{k}\alpha}a_{\mathbf{k}\alpha} \tag{2.70}$$

and $H_F = \sum_{\mathbf{k}\alpha}\hbar\omega_k a^\dagger_{\mathbf{k}\alpha}a_{\mathbf{k}\alpha}, \mathbf{P} = \sum_{\mathbf{k}\alpha}\hbar\mathbf{k}a^\dagger_{\mathbf{k}\alpha}a_{\mathbf{k}\alpha}$. With circularly polarized mode functions, therefore, $\mathbf{M}_s$ commutes with $H_F$ and has the photon number state $|n_{\mathbf{k}\alpha}\rangle$ as an eigenstate with eigenvalue $\alpha\hbar\hat{\mathbf{k}}, \alpha = \pm 1$. In other words, the component of the photon spin along the direction of propagation, the photon *helicity*, is $\pm 1$ in units of $\hbar$, which means that a photon is a boson of spin 1. Ordinarily we will not be concerned with spin and will employ the linear polarization basis.

## 2.6 Necessity of the Vacuum Field

The vacuum state $|\text{vac}\rangle$ of the free field is defined as the ground state in which $n_{\mathbf{k}\lambda} = 0$ for all modes $(\mathbf{k}, \lambda)$. The vacuum state, like all stationary states of the field, is an eigenstate of the Hamiltonian but not the electric and magnetic field operators. In the vacuum state, therefore, the electric and magnetic fields do not have definite values. We can imagine them to be fluctuating about their mean values of zero, as discussed in Section 2.4 for the case of a single mode of the field.

In a process in which a photon is annihilated (absorbed), we can think of the photon as making a transition into the vacuum state. Similarly, when a photon is created (emitted), it is occasionally useful to imagine that the photon has made a transition out of the vacuum state. In the words of Dirac (1927),

> The light-quantum has the peculiarity that it apparently ceases to exist when it is in one of its stationary states, namely, the zero state, in which its momentum, and therefore also its energy, are zero. When a light-quantum is absorbed it can be considered to jump into this zero state, and when one is emitted it can be considered to jump from the zero state to one in which it is physically in evidence, so that it appears to have been created. Since there is no limit to the number of light-quanta that may be created in this way, we must suppose that there are an infinite number of light-quanta in the zero state ...



We shall see later that an atom, for instance, can be considered to be "dressed" by emission and reabsorption of "virtual photons" from the vacuum.

The most glaring characteristic of the vacuum state is that its energy $\sum_{\mathbf{k}\lambda}\frac{1}{2}\hbar\omega_k$ is infinite. Let us use (2.47) to make the well-known replacement

$$\sum_{\mathbf{k}\lambda} \to \sum_{\lambda}\left(\frac{L}{2\pi}\right)^3 \int d^3k = \frac{V}{8\pi^3}\sum_{\lambda}\int d^3k. \tag{2.71}$$

The zero-point energy density is thus

$$\begin{aligned}\frac{1}{V}\sum_{\mathbf{k}\lambda}\frac{1}{2}\hbar\omega_k &= \frac{2}{8\pi^3}\int d^3k\frac{1}{2}\hbar\omega_k = \frac{4\pi}{4\pi^3}\int dk k^2(\frac{1}{2}\hbar\omega_k) \\ &= \frac{\hbar}{2\pi^2c^3}\int d\omega\omega^3,\end{aligned} \tag{2.72}$$

or in other words the spectral energy density of the vacuum field is

$$\rho_o(\omega) = \frac{\hbar\omega^3}{2\pi^2c^3}, \tag{2.73}$$

which is familiar from Chapter 1. The zero-point energy density in the frequency range from $\omega_1$ to $\omega_2$ is therefore

$$\int_{\omega_1}^{\omega_2} d\omega\rho_o(\omega) = \frac{\hbar}{8\pi^2c^3}(\omega_2^4 - \omega_1^4). \tag{2.74}$$

This can be large even in relatively narrow, "low-frequency" regions of the spectrum. In the optical region from 400 nm to 700 nm, for instance, equation (2.74) yields about 220 erg/cm$^3$.

In Section 2.4 we noted that the zero-point energy of the field can be eliminated from the Hamiltonian by the normal ordering prescription. However, this elimination does not mean that the vacuum field has been rendered unimportant or without physical consequences! To illustrate this point we consider now a linear dipole oscillator in the vacuum.

The Hamiltonian for the oscillator plus the field with which it interacts is

$$H = \frac{1}{2m}(\mathbf{p} - \frac{e}{c}\mathbf{A})^2 + \frac{1}{2}m\omega_o^2\mathbf{x}^2 + H_F. \tag{2.75}$$

Of course this has the same form as the corresponding classical Hamiltonian, and the Heisenberg equations of motion for the oscillator and the field are formally the same as their classical counterparts. For instance, the Heisenberg equations for the coordinate $\mathbf{x}$ and the canonical momentum $\mathbf{p} = m\dot{\mathbf{x}} + e\mathbf{A}/c$ of the oscillator are[2]

$$\dot{\mathbf{x}} = (i\hbar)^{-1}[\mathbf{x}, H] = \frac{1}{m}(\mathbf{p} - \frac{e}{c}\mathbf{A}), \tag{2.76}$$

$$\begin{aligned}\dot{\mathbf{p}} = (i\hbar)^{-1}[\mathbf{p}, H] &= -\frac{1}{2m}\nabla(\mathbf{p} - \frac{e}{c}\mathbf{A})^2 - m\omega_o^2\mathbf{x} \\ &= -\frac{1}{m}[(\mathbf{p} - \frac{e}{c}\mathbf{A}) \cdot \nabla][-\frac{e}{c}\mathbf{A}] - \frac{1}{m}(\mathbf{p} - \frac{e}{c}\mathbf{A}) \\ &\qquad \times \nabla \times [-\frac{e}{c}\mathbf{A}] - m\omega_o^2\mathbf{x} \\ &= \frac{e}{c}(\dot{\mathbf{x}} \cdot \nabla)\mathbf{A} + \frac{e}{c}\dot{\mathbf{x}} \times \mathbf{B} - m\omega_o^2\mathbf{x},\end{aligned} \tag{2.77}$$

or

$$\begin{aligned}m\ddot{\mathbf{x}} &= \dot{\mathbf{p}} - \frac{e}{c}\dot{\mathbf{A}} = -\frac{e}{c}[\dot{\mathbf{A}} - (\dot{\mathbf{x}} \cdot \nabla)\mathbf{A}] + \frac{e}{c}\dot{\mathbf{x}} \times \mathbf{B} - m\omega_o^2\mathbf{x} \\ &= e\mathbf{E} + \frac{e}{c}\dot{\mathbf{x}} \times \mathbf{B} - m\omega_o^2\mathbf{x},\end{aligned} \tag{2.78}$$

since the rate of change of the vector potential in the frame of the moving charge is given by the convective derivative $\dot{\mathbf{A}} = \partial\mathbf{A}/\partial t + (\dot{\mathbf{x}} \cdot \nabla)\mathbf{A}$.[3] For nonrelativistic motion we may neglect the magnetic force and replace (2.78) by

$$\ddot{\mathbf{x}} + \omega_o^2\mathbf{x} \cong \frac{e}{m}\mathbf{E} \cong i\frac{e}{m}\sum_{\mathbf{k}\lambda}\left(\frac{2\pi\hbar\omega_k}{V}\right)^{1/2}[a_{\mathbf{k}\lambda}(t) - a_{\mathbf{k}\lambda}^\dagger(t)]\mathbf{e}_{\mathbf{k}\lambda}. \tag{2.79}$$

As in Chapter 1 we have made the electric dipole approximation in which the spatial dependence of the field is neglected. The Heisenberg equation

[2]The Hamiltonian for a charged particle in an electromagnetic field is reviewed in Chapter 4. In (2.77) we employ the vector generalization of the identity $[p, F(q,p)] = -i\hbar\partial F/\partial q$.

[3]This follows from the general relation $i\hbar dA_x/dt = [A_x, H] + i\hbar\partial A_x/\partial t$.



for $a_{\mathbf{k}\lambda}$ is found similarly from the Hamiltonian (2.75) to be

$$\dot{a}_{\mathbf{k}\lambda} = -i\omega_k a_{\mathbf{k}\lambda} + ie\left(\frac{2\pi}{\hbar\omega_k V}\right)^{1/2}\dot{\mathbf{x}} \cdot \mathbf{e}_{\mathbf{k}\lambda} \tag{2.80}$$

in the electric dipole approximation. In deriving these equations for $\mathbf{x}$, $\mathbf{p}$, and $a_{\mathbf{k}\lambda}$ we have used the fact that equal-time particle and field operators commute. This follows from the assumption that particle and field operators commute at some time (say, $t = 0$) when the matter–field interaction is presumed to begin, together with the fact that a Heisenberg-picture operator $A(t)$ evolves in time as $A(t) = U^\dagger(t)A(0)U(t)$, where $U(t)$ is the time evolution operator satisfying $i\hbar\dot{U} = HU$, $U^\dagger(t) = U^{-1}(t)$, $U(0) = 1$. Alternatively, we can argue that these operators must commute if we are to obtain the correct equations of motion from the Hamiltonian, just as the corresponding Poisson brackets in classical theory must vanish in order to generate the correct Hamilton equations (see also Section 4.2).

The formal solution of the field equation (2.80) is

$$a_{\mathbf{k}\lambda}(t) = a_{\mathbf{k}\lambda}(0)e^{-i\omega_k t} + ie\left(\frac{2\pi}{\hbar\omega_k V}\right)^{1/2}\int_0^t dt'\mathbf{e}_{\mathbf{k}\lambda} \cdot \dot{\mathbf{x}}(t')e^{i\omega_k(t'-t)}, \tag{2.81}$$

and therefore equation (2.79) may be written

$$\ddot{\mathbf{x}} + \omega_o^2\mathbf{x} = \frac{e}{m}\mathbf{E}_o(t) + \frac{e}{m}\mathbf{E}_{RR}(t), \tag{2.82}$$

where

$$\mathbf{E}_o(t) = i\sum_{\mathbf{k}\lambda}\left(\frac{2\pi\hbar\omega_k}{V}\right)^{1/2}[a_{\mathbf{k}\lambda}(0)e^{-i\omega_k t} - a_{\mathbf{k}\lambda}^\dagger(0)e^{i\omega_k t}]\mathbf{e}_{\mathbf{k}\lambda} \tag{2.83}$$

and

$$\mathbf{E}_{RR}(t) = -\frac{4\pi e}{V}\sum_{\mathbf{k}\lambda}\int_0^t dt'[\mathbf{e}_{\mathbf{k}\lambda} \cdot \dot{\mathbf{x}}(t')]\mathbf{e}_{\mathbf{k}\lambda}\cos\omega_k(t'-t). \tag{2.84}$$

We show in Appendix D that we may take

$$\mathbf{E}_{RR}(t) = \frac{2e}{3c^3}\dddot{\mathbf{x}} \tag{2.85}$$

for the radiation reaction field, if the mass $m$ in (2.82) is regarded as the "observed" mass.

The total field acting on the dipole has two parts, $\mathbf{E}_o(t)$ and $\mathbf{E}_{RR}(t)$. $\mathbf{E}_o(t)$ is the free or zero-point field acting on the dipole. It is the homogeneous solution of the Maxwell equation for the field acting on the dipole, i.e., the solution, at the position of the dipole, of the wave equation $[\nabla^2 - c^{-2}\partial^2/\partial t^2]\mathbf{E} = 0$ satisfied by the field in the (source-free) vacuum. For this reason $\mathbf{E}_o(t)$ is often referred to as the *vacuum field*, although it is of course a Heisenberg-picture operator acting on whatever state of the field happens to be appropriate at $t = 0$. $\mathbf{E}_{RR}(t)$ is the source field, the field generated *by* the dipole and acting *on* the dipole.

Using (2.85) in (2.82), we obtain an equation for the Heisenberg-picture operator $\mathbf{x}(t)$ that is formally the same as the classical equation (1.41):

$$\ddot{\mathbf{x}} + \omega_o^2\mathbf{x} - \tau\dddot{\mathbf{x}} = \frac{e}{m}\mathbf{E}_o(t), \tag{2.86}$$

where again $\tau = 2e^2/3mc^3$. But here we have considered a dipole in the vacuum, without any "external" field acting on it. The role of the "external" field in equation (2.86) is played by the *vacuum* electric field acting on the dipole.

Classically, of course, a dipole in the vacuum is not acted upon by any "external" field: if there are no sources other than the dipole itself, then the only field acting on the dipole is its own radiation reaction field. In quantum theory, however, there is always an "external" field, namely, the source-free or vacuum field $\mathbf{E}_o(t)$.

According to equation (2.81) the free field is the only field in existence at $t = 0$. This defines $t = 0$ as the time at which the interaction between the dipole and the field is "switched on." The state vector of the dipole–field system at $t = 0$ is therefore of the form $|\Psi\rangle = |\text{vac}\rangle|\psi_D\rangle$, where $|\text{vac}\rangle$ is the vacuum state of the field and $|\psi_D\rangle$ is the initial state of the dipole oscillator. The expectation value of the free field is therefore at all times equal to zero: $\langle\mathbf{E}_o(t)\rangle = \langle\Psi|\mathbf{E}_o(t)|\Psi\rangle = 0$ since $a_{\mathbf{k}\lambda}(0)|\text{vac}\rangle = 0$. However, the energy density associated with the free field is infinite:

$$\begin{aligned}\frac{1}{4\pi}\langle\mathbf{E}_o^2(t)\rangle &= \frac{1}{4\pi}\sum_{\mathbf{k}\lambda}\sum_{\mathbf{k}'\lambda'}\left(\frac{2\pi\hbar\omega_k}{V}\right)^{1/2}\left(\frac{2\pi\hbar\omega_{k'}}{V}\right)^{1/2} \\ &\quad\times\langle a_{\mathbf{k}\lambda}(0)a^\dagger_{\mathbf{k}'\lambda'}(0)\rangle \\ &= \frac{1}{4\pi}\sum_{\mathbf{k}\lambda}\left(\frac{2\pi\hbar\omega_k}{V}\right) = \int_0^\infty d\omega\rho_o(\omega).\end{aligned} \tag{2.87}$$

The important point is this: the zero-point field energy in $H_F$ does not affect the Heisenberg equation for $a_{\mathbf{k}\lambda}$, since it is a $c$-number (i.e., an ordinary number rather than an operator) and commutes with $a_{\mathbf{k}\lambda}$. We can therefore drop the zero-point field energy from the Hamiltonian, as is usually done. But the zero-point field re-emerges, so to speak, as the homogeneous solution of the field equation. *A charged particle in the vacuum will therefore always see a zero-point field of infinite energy density.* This is the origin of one of the infinities of quantum electrodynamics, and it cannot be eliminated by the trivial expedient of dropping the term $\sum_{\mathbf{k}\lambda}\frac{1}{2}\hbar\omega_k$ in the field Hamiltonian.



The free field is in fact *necessary* for the formal consistency of the theory. In particular, it is necessary for the preservation of commutation relations, which is required by the unitarity of time evolution in quantum theory: $[z(t), p_z(t)] = [U^\dagger(t)z(0)U(t), U^\dagger(t)p_z(0)U(t)] = U^\dagger(t)[z(0), p_z(0)]U(t) = i\hbar U^\dagger(t)U(t) = i\hbar$. We can calculate $[z(t), p_z(t)]$ from the formal solution of the operator equation of motion (2.86). Using the fact that $[a_{\mathbf{k}\lambda}(0), a^\dagger_{\mathbf{k}'\lambda'}(0)] = \delta^3_{\mathbf{k},\mathbf{k}'}\delta_{\lambda\lambda'}$, and that equal-time particle and field operators commute, we readily obtain

$$\begin{aligned}[z(t), p_z(t)] &= [z(t), m\dot{z}(t)] + [z(t), \frac{e}{c}A_z(t)] = [z(t), m\dot{z}(t)] \\ &= \left(\frac{i\hbar e^2}{2\pi^2 mc^3}\right)\left(\frac{8\pi}{3}\right)\int_0^\infty \frac{d\omega\omega^4}{(\omega^2-\omega_o^2)^2+\tau^2\omega^6}\end{aligned} \tag{2.88}$$

in the mode continuum limit (2.71). For the dipole oscillator under consideration it can sensibly be assumed that the radiative damping rate is small compared with the natural oscillation frequency, i.e., $\tau\omega_o << 1$. Then the integrand in (2.88) is sharply peaked at $\omega = \omega_o$, and[4]

$$\begin{aligned}[z(t), p_z(t)] &\cong \frac{2i\hbar e^2}{3\pi mc^3}\omega_o^3\int_{-\infty}^\infty \frac{dx}{x^2+\tau^2\omega_o^6} = \left(\frac{2i\hbar e^2\omega_o^3}{3\pi mc^3}\right)\left(\frac{\pi}{\tau\omega_o^3}\right) \\ &= i\hbar.\end{aligned} \tag{2.89}$$

We can appreciate further the necessity of the vacuum field by making the small-damping approximation directly in (2.86): $\ddot{\mathbf{x}} \cong -\omega_o^2\mathbf{x}(t), \dddot{\mathbf{x}} \cong -\omega_o^2\dot{\mathbf{x}}$, and

$$\ddot{\mathbf{x}} + \tau\omega_o^2\dot{\mathbf{x}} + \omega_o^2\mathbf{x} \cong \frac{e}{m}\mathbf{E}_o(t). \tag{2.90}$$

Without the free field $\mathbf{E}_o(t)$ in this equation the *operator* $\mathbf{x}(t)$ would be exponentially damped, and commutators like $[z(t), p_z(t)]$ would approach zero for $t >> (\tau\omega_o^2)^{-1}$. With the vacuum field included, however, the

[4] Actually (2.89) follows *exactly* from (2.88), as may be shown using the residue theorem.

commutator is $i\hbar$ at all times, as required by unitarity, and as we have just shown. A similar result is easily worked out for the case of a free particle instead of a dipole oscillator (Milonni, 1981b).

What we have here is an example of a "fluctuation–dissipation relation." Generally speaking, if a system is coupled to a "bath" that can take energy from the system in an effectively irreversible way, then the bath must also cause fluctuations. The fluctuations and the dissipation go hand in hand; we cannot have one without the other. In the present example the coupling of a dipole oscillator to the electromagnetic field has a dissipative component, in the form of radiation reaction, and a fluctuation component, in the form of the zero-point (vacuum) field; given the existence of radiation reaction, the vacuum field must also exist in order to preserve the canonical commutation rule and all it entails.

The spectral density of the vacuum field is fixed by the form of the radiation reaction field, or vice versa: because the radiation reaction field varies with the *third* derivative of $\mathbf{x}$, the spectral energy density of the vacuum field must be proportional to the *third* power of $\omega$ in order for (2.88) to hold. In the case of a dissipative force proportional to $\dot{\mathbf{x}}$, by contrast, the fluctuation force must be proportional to $\omega$ in order to maintain the canonical commutation relation (Milonni, 1981b). This relation between the form of the dissipation and the spectral density of the fluctuation is the essence of the fluctuation–dissipation theorem.[5]

The fact that the canonical commutation relation for a harmonic oscillator coupled to the vacuum field is preserved implies that the zero-point energy of the oscillator is preserved. It is easy to show that after a few damping times the zero-point motion of the oscillator is in fact sustained by the driving zero-point field (Senitzky, 1960).

The reader may well wonder whether the vacuum field is merely some sort of formal mathematical artifice of quantum electrodynamics, whether it really has any unambiguous experimental manifestations. In fact the zero-point field does appear to be quite "real," as we shall see in the following section.

## 2.7 The Casimir Effect

Casimir showed in 1948 that one consequence of the zero-point field is an attractive force between two uncharged, perfectly conducting parallel plates (Figure 2.1). In this section we review a standard calculation of the Casimir force, and in the following chapter we present a somewhat more physical

[5] H. B. Callen and T. A. Welton, *Phys. Rev.* **83**, 34 (1951).



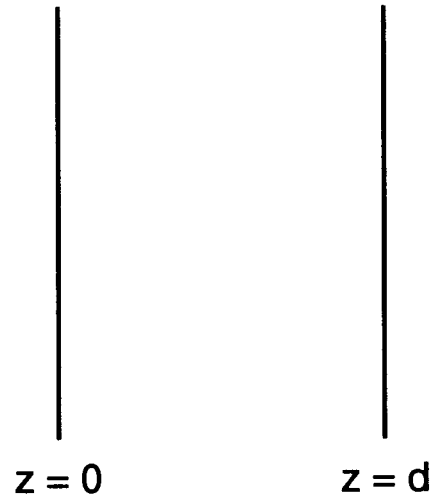


Figure 2.1: Two conducting parallel plates experience an attractive force attributable to the zero-point electromagnetic field. This is the Casimir effect.

variation of this calculation. Various Casimir effects, and experimental evidence for them, are discussed in Chapters 7 and 8.

The physical situation shown in Figure 2.1 leads us to consider a different set of modes than the free-space plane-wave modes we have dealt with thus far. Consider first the modes appropriate to the interior of a rectangular parallelepiped of sides $L_x = L_y = L$ and $L_z$. For perfectly conducting walls the mode functions satisfying the boundary condition that the tangential component of the electric field vanishes on the walls are $\mathbf{A}(\mathbf{r}) = A_x(\mathbf{r})\mathbf{i} + A_y(\mathbf{r})\mathbf{j} + A_z(\mathbf{r})\mathbf{k}$, where

$$A_x(\mathbf{r}) = (8/V)^{1/2} a_x \cos(k_x x)\sin(k_y y)\sin(k_z z), \tag{2.91}$$

$$A_y(\mathbf{r}) = (8/V)^{1/2} a_y \sin(k_x x)\cos(k_y y)\sin(k_z z), \tag{2.92}$$

$$A_z(\mathbf{r}) = (8/V)^{1/2} a_z \sin(k_x x)\sin(k_y y)\cos(k_z z), \tag{2.93}$$

with $a_x^2 + a_y^2 + a_z^2 = 1, V = L^2 L_z$, and

$$k_x = \frac{\ell\pi}{L}, \quad k_y = \frac{m\pi}{L}, \quad k_z = \frac{n\pi}{L_z}, \tag{2.94}$$

with $\ell$, $m$, and $n$ each taking on all positive integer values and zero. In

order to satisfy the transversality condition $\nabla \cdot \mathbf{A} = 0$ we also require

$$k_x A_x + k_y A_y + k_z A_z = \frac{\pi}{L}(\ell A_x + m A_y) + \frac{\pi}{L_z}(n A_z) = 0. \tag{2.95}$$

Thus there are two independent polarizations, unless one of the integers $\ell$, $m$, or $n$ is zero, in which case (2.95) indicates that there is only one polarization. It is easy to check that equations (2.91)–(2.93) define transverse mode functions satisfying the Helmholtz equation (2.24) as well as the condition that the transverse components of $\mathbf{E}$ vanish on the cavity walls. Furthermore these mode functions are orthogonal and satisfy the normalization condition (2.30), i.e.,

$$\int_0^L dx \int_0^L dy \int_0^{L_z} dz [A_x^2(\mathbf{r}) + A_y^2(\mathbf{r}) + A_z^2(\mathbf{r})] = 1. \tag{2.96}$$

Actually all we really require for the calculation of the Casimir force are the allowed frequencies defined by (2.94):

$$\omega_{\ell mn} = k_{\ell mn} c = \pi c \left[\frac{\ell^2}{L^2} + \frac{m^2}{L^2} + \frac{n^2}{L_z^2}\right]^{1/2}. \tag{2.97}$$

The zero-point energy of the field inside the cavity is therefore

$$\sum_{\ell,m,n}{}'(2)\frac{1}{2}\hbar\omega_{\ell mn} = \sum_{\ell mn}{}'\pi\hbar c \left[\frac{\ell^2}{L^2} + \frac{m^2}{L^2} + \frac{n^2}{L_z^2}\right]^{1/2}. \tag{2.98}$$

The factor 2 arises from the two independent polarizations of modes with $\ell, m, n \neq 0$, and the prime on the summation symbol implies that a factor $1/2$ should be inserted if one of these integers is zero, for then we have just one independent polarization, as noted earlier.

In the physical situation of interest $L$ is so large compared with $L_z = d$ that we may replace the sums over $\ell$ and $m$ in (2.98) by integrals: $\sum_{\ell mn} \to \sum_n{}'(L/\pi)^2 \int\int dk_x dk_y$ and

$$E(d) = \sum_{\ell mn}{}'(2)\frac{1}{2}\hbar\omega_{\ell mn} \to \frac{L^2}{\pi^2}(\hbar c)\sum_n{}' \int_0^\infty dk_x \int_0^\infty dk_y \times \left(k_x^2 + k_y^2 + \frac{n^2\pi^2}{d^2}\right)^{1/2}. \tag{2.99}$$

This is infinite; the zero-point energy of the vacuum is infinite in *any* finite volume.



If $d$ were also made arbitrarily large, the sum over $n$ could be replaced by an integral. Then the zero-point energy (2.99) would be

$$E(\infty) = \frac{L^2}{\pi^2}(\hbar c)\frac{d}{\pi}\int_0^\infty dk_x \int_0^\infty dk_y \int_0^\infty dk_z (k_x^2 + k_y^2 + k_z^2)^{1/2}. \tag{2.100}$$

which is also infinite.

The potential energy of the system when the plates are separated by a distance $d$ is $U(d) = E(d) - E(\infty)$, the energy required to bring the plates from a large separation to the separation $d$:

$$U(d) = \frac{L^2\hbar c}{\pi^2}\left[\sum_n{}' \int_0^\infty dk_x \int_0^\infty dk_y (k_x^2 + k_y^2 + \frac{n^2\pi^2}{d^2})^{\frac{1}{2}} - \frac{d}{\pi}\int_0^\infty dk_x \int_0^\infty dk_y \int_0^\infty dk_z (k_x^2 + k_y^2 + k_z^2)^{1/2}\right]. \tag{2.101}$$

This is the difference between two infinite quantities, but we shall now show that it is nonetheless possible to extract from it a physically meaningful, finite value.[6]

In polar coordinates $u, \theta$ in the $k_x, k_y$ plane ($dk_x dk_y = u du d\theta$) we have

$$U(d) = \frac{L^2\hbar c}{\pi^2}\left(\frac{\pi}{2}\right)\left[\sum_{n=0}{}' \int_0^\infty du u \left(u^2 + \frac{n^2\pi^2}{d^2}\right)^{1/2} - \left(\frac{d}{\pi}\right)\int_0^\infty dk_z \int_0^\infty du u (u^2 + k_z^2)^{1/2}\right], \tag{2.102}$$

since $\theta$ ranges from 0 to $\pi/2$ for $k_x, k_y > 0$. We now introduce a cutoff function $f(k) = f([u^2 + k_z^2]^{1/2})$ such that $f(k) = 1$ for $k << k_m$ and $f(k) = 0$ for $k >> k_m$. Physically, it can be argued that $f(k)$ is necessary because the assumption of perfectly conducting walls breaks down at small wavelengths and especially for wavelengths small compared with an atomic dimension. We might then suppose that $k_m \approx 1/a_o$, where $a_o$ is the Bohr radius. What we are assuming here is that the Casimir effect is primarily a low-frequency, nonrelativistic effect. We thus replace (2.102) by

$$U(d) = \frac{L^2\hbar c}{\pi^2}\left(\frac{\pi}{2}\right)\left[\sum_{n=0}{}' \int_0^\infty du u (u^2 + \frac{n^2\pi^2}{d^2})^{1/2} f([u^2 + \frac{n^2\pi^2}{d^2}]^{1/2}) - \left(\frac{d}{\pi}\right)\int_0^\infty dk_z \int_0^\infty du u (u^2 + k_z^2)^{1/2} f([u^2 + k_z^2]^{1/2})\right]$$

[6] See Section 10.7 for a different approach.

$$= \frac{L^2\hbar c}{4\pi}\left(\frac{\pi^3}{d^3}\right)\left[\sum_{n=0}{}' \int_0^\infty dx(x+n^2)^{1/2} f\left(\frac{\pi}{d}[x+n^2]^{1/2}\right) - \int_0^\infty d\kappa \int_0^\infty dx(x+\kappa^2)^{1/2} f\left(\frac{\pi}{d}[x+\kappa^2]^{1/2}\right)\right], \quad (2.103)$$

where we have defined the new integration variables $x = u^2d^2/\pi^2$ and $\kappa = k_z d/\pi$. Now

$$U(d) = \left(\frac{\pi^2\hbar c}{4d^3}\right) L^2 \left[\frac{1}{2}F(0) + \sum_{n=1}^\infty F(n) - \int_0^\infty d\kappa F(\kappa)\right], \quad (2.104)$$

where

$$F(\kappa) \equiv \int_0^\infty dx(x+\kappa^2)^{1/2} f\left(\frac{\pi}{d}[x+\kappa^2]^{1/2}\right). \quad (2.105)$$

According to the Euler–Maclaurin summation formula[7]

$$\sum_{n=1}^\infty F(n) - \int_0^\infty d\kappa F(\kappa) = -\frac{1}{2}F(0) - \frac{1}{12}F'(0) + \frac{1}{720}F'''(0) \ldots \quad (2.106)$$

for $F(\infty) \to 0$. To evaluate the nth derivative $F^{(n)}(0)$ we note that

$$F(\kappa) = \int_{\kappa^2}^\infty du\sqrt{u} f\left(\frac{\pi}{d}\sqrt{u}\right), \quad F'(\kappa) = -2\kappa^2 f\left(\frac{\pi}{d}\kappa\right). \quad (2.107)$$

Then $F'(0) = 0$, $F'''(0) = -4$, and all higher derivatives $F^{(n)}(0)$ vanish if we assume that all derivatives of the cutoff function vanish at $\kappa = 0$. Thus $\sum_{n=1}^\infty F(n) - \int_0^\infty d\kappa F(\kappa) = -\frac{1}{2}F(0) - \frac{4}{720}$ and

$$U(d) = \left(\frac{\pi^2\hbar c}{4d^3}\right) L^2 \left(\frac{-4}{720}\right) = -\left(\frac{\pi^2\hbar c}{720d^3}\right) L^2, \quad (2.108)$$

which is finite and independent of the cutoff function. The attractive force per unit area between the plates is then $F(d) = -\pi^2\hbar c/240d^4$. This is the Casimir force, which we shall revisit in the following chapter and again in Chapters 7 and 8. The principal message of this section is that *changes* in the infinite zero-point energy of the electromagnetic vacuum can be finite and observable.

[7] See M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover Books, New York, 1971), Formula 3.6.28. For a derivation of the Euler–Maclaurin formula, see, for instance, E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, 4th ed. (Cambridge University Press, New York, 1969), p. 127.



## 2.8 Field Commutators

The fundamental field commutator (2.55) holds for all times $t$, and regardless of whether there are any sources of radiation. From this commutator one readily obtains commutation relations for the field vectors, such as

$$\begin{aligned}[E_i(\mathbf{r}_1,t), E_j(\mathbf{r}_2,t_2)] &= [B_i(\mathbf{r}_1,t), B_j(\mathbf{r}_2,t_2)] \\ &= 4\pi i\hbar c\left(\frac{\delta_{ij}}{c^2}\frac{\partial^2}{\partial t_1 \partial t_2} - \frac{\partial^2}{\partial r_{1i}\partial r_{2j}}\right) \\ &\quad \times D(|\mathbf{r}_1 - \mathbf{r}_2|, t_1 - t_2), \end{aligned} \quad (2.109)$$

where

$$\begin{aligned} D(\mathbf{r},t) &\equiv -\left(\frac{1}{2\pi}\right)^3 \int d^3k \frac{1}{k} e^{i\mathbf{k}\cdot\mathbf{r}} \sin\omega_k t \\ &= -\frac{1}{2\pi^2 r}\int_{-\infty}^\infty dk \sin kr \sin kct \\ &= \frac{1}{4\pi r}[\delta(r+ct) - \delta(r-ct)]. \end{aligned} \quad (2.110)$$

These "Pauli–Jordan commutators" imply that the fields at space–time points $(\mathbf{r}_1, t_1)$ and $(\mathbf{r}_2, t_2)$ cannot in general be simultaneously measured if these points can be connected by a light signal, i.e., if $|\mathbf{r}_1 - \mathbf{r}_2| = \pm c(t_1 - t_2)$. Similarly

$$[E_i(\mathbf{r}_1,t_1), B_j(\mathbf{r}_2,t_2)] = 4\pi i\hbar c\epsilon_{ijk}\frac{\delta^2}{\delta t_1 \delta r_{2k}} D(|\mathbf{r}_1 - \mathbf{r}_2|, t_1 - t_2). \quad (2.111)$$

The physical significance of these commutators was discussed by Bohr and Rosenfeld (1950): since the field of a charged particle provides information about the motion of the particle, the uncertainty relations ($\Delta x \Delta p_x \geq \hbar/2$, etc.) for the particles must, for the consistency of quantum theory, imply uncertainty relations also for the field. These uncertainty relations for the electromagnetic field are embodied in the Pauli–Jordan commutators. Note that $D(\mathbf{r},t) = -D(\mathbf{r},-t)$, and so $\lim_{t\to 0} D(\mathbf{r},t) = \lim_{t\to 0}(\partial^2/\partial t^2)D(\mathbf{r},t) = 0$. Equation (2.110) then indicates that in principle the electric and magnetic fields can be simultaneously measured everywhere in space at a fixed instant of time.

Note also that these field commutators are derived for free space. The presence of boundaries or even simple point sources will in general lead to different commutation relations, simply because $e^{i\mathbf{k}\cdot\mathbf{r}}$ in (2.110) must be replaced by different mode functions.[8]

[8] This is discussed in the papers by Milonni (1982) and Cresser (1984) cited at the end of the chapter.

## 2.9 Zero-Point Spectrum Invariance

We have seen in connection with the classical Einstein–Hopf model that a dipole oscillating with frequency $\omega$, moving with velocity $v$ through a thermal field, experiences a frictional force $F = -Rv$, where $R$ is proportional to $\rho(\omega) - (\omega/3)d\rho/d\omega$. The same result holds when the dipole is treated quantum mechanically or if, as shown in Appendix B, the dipole is replaced by an atom.

At $T = 0$ we have $\rho(\omega) = \rho_0(\omega) = \hbar\omega^3/2\pi^2c^3$ and

$$\rho_0(\omega) - \frac{\omega}{3}\frac{d\rho_0}{d\omega} = 0. \tag{2.112}$$

In other words, there is no frictional force acting on a dipole or atom moving with constant velocity in the vacuum. The zero-point spectrum proportional to $\omega^3$, which is precisely the form required by the fluctuation–dissipation relation (Section 2.6), is thus the the *unique* spectral energy density for which there is no force. Alternatively, we can say that, since the number of modes per unit volume in free space is proportional to $\omega^2$, the energy $\frac{1}{2}\hbar\omega$ per mode is the unique zero-point energy for which there is no force. The "uniqueness" refers, of course, to the functional dependence on $\omega$; any zero-point energy proportional to $\omega$, or any spectral energy density proportional to $\omega^3$, will satisfy (2.112).

In fact it has been shown explicitly by Boyer (1969), using the Lorentz transformations for the electric and magnetic fields, that $\rho_0(\omega)$ is the unique Lorentz-invariant spectral energy density of the electromagnetic field. That is, the condition that $\rho_0(\omega)$ be the same in all inertial frames requires it to be proportional to $\omega^3$. This conforms with our expectation that an observer moving with constant velocity in the electromagnetic vacuum cannot tell that he is moving!

## 2.10 The Unruh–Davies Effect

What if the observer is moving with constant (proper) *acceleration* in the vacuum? Then a remarkable thing happens: the observer perceives himself to be immersed in a thermal bath at the temperature $T = \hbar a/2\pi kc$, where $a$ is the acceleration. This result was obtained by Unruh (1976), following a closely related result of Davies (1975). In this section we shall demonstrate this thermal effect of acceleration for the case of a scalar field, for which the calculation is simpler. The electromagnetic case is somewhat more complicated but the result is the same.



We consider a massless scalar field $\phi(\mathbf{x},t)$ satisfying the wave equation $(\nabla^2 - c^{-2}\partial^2/\partial t^2)\phi = 0$ and having an energy density $(1/8\pi)[(\nabla\phi)^2 + c^{-2}(\partial\phi/\partial t)^2]$.[9] When quantized in free space $\phi(\mathbf{x},t)$ has the form

$$\phi(\mathbf{x},t) = \sum_{\mathbf{k}} \left(\frac{2\pi\hbar c^2}{\omega_k V}\right)^{1/2} [a_{\mathbf{k}}(t)e^{i\mathbf{k}\cdot\mathbf{x}} + a_{\mathbf{k}}^\dagger(t)e^{-i\mathbf{k}\cdot\mathbf{x}}], \tag{2.113}$$

where again we assume periodic boundary conditions. Here, $a_{\mathbf{k}}(t)$ and $a_{\mathbf{k}}^\dagger(t)$ are boson annihilation and creation operators and $a_{\mathbf{k}}(t) = a_{\mathbf{k}}(0)e^{-i\omega_k t}$ for the free field, with $\omega_k = kc$. The Hamiltonian is

$$H_{\mathrm{F}} = \frac{1}{8\pi}\int d^3x \left[(\nabla\phi)^2 + \frac{1}{c^2}(\frac{\partial\phi}{\partial t})^2\right] = \sum_{\mathbf{k}} \hbar\omega_k(a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \frac{1}{2}). \tag{2.114}$$

Everything here is much the same as in the case of the electromagnetic field, but simpler.

Consider the field correlation function $\langle\phi(0,t)\phi(0,t+\tau)\rangle$ at a point in space for a field in thermal equilibrium at temperature $T$. In this case $\langle a_{\mathbf{k}}^\dagger(0)a_{\mathbf{k}'}(0)\rangle = \delta^3_{\mathbf{k},\mathbf{k}'}\overline{n}(\omega)$, $\overline{n}(\omega) \equiv (e^{\hbar\omega/kT} - 1)^{-1}$. These results are intuitively obvious. Basically they imply that different modes of a thermal field are uncorrelated, each mode amplitude having zero expectation value, and that a mode of frequency $\omega$ has an average number of quanta $\overline{n}(\omega)$. Thus

$$\begin{aligned}
\langle\phi(0,t)\phi(0,t+\tau)\rangle &= \sum_{\mathbf{k}} \left(\frac{2\pi\hbar c^2}{\omega_k V}\right)\left[\langle a_{\mathbf{k}}(t)a_{\mathbf{k}}^\dagger(t+\tau)\rangle\right. \\
&\quad \left. + \langle a_{\mathbf{k}}^\dagger(t)a_{\mathbf{k}}(t+\tau)\rangle\right] \\
&= \sum_{\mathbf{k}} \left(\frac{2\pi\hbar c^2}{\omega_k V}\right)\left[(\overline{n}(\omega_k)+1)e^{i\omega_k\tau} + \overline{n}(\omega_k)e^{-i\omega_k\tau}\right] \\
&\to \frac{\hbar}{\pi c}\left[\int_0^\infty d\omega\omega e^{i\omega\tau} + 2\int_0^\infty \frac{d\omega\omega\cos\omega\tau}{e^{\hbar\omega/kT}-1}\right].
\end{aligned} \tag{2.115}$$

The first integral may be evaluated as follows:

$$\int_0^\infty d\omega\omega e^{i\omega\tau} = \lim_{x\to 0}\int_0^\infty d\omega\omega e^{i\omega(\tau+ix)} = \lim_{x\to 0}\frac{1}{(x-i\tau)^2} = -\frac{1}{\tau^2}. \tag{2.116}$$

[9] See Section 10.3.

The second integral follows from the general formula[10]

$$\int_0^\infty dx \frac{x^{2m+1}\cos bx}{e^x-1} = (-1)^m \frac{\partial^{2m+1}}{\partial b^{2m+1}}\left[\frac{\pi}{2}\coth \pi b - \frac{1}{2b}\right]. \tag{2.117}$$

Thus

$$\int_0^\infty \frac{d\omega\omega\cos\omega\tau}{e^{\hbar\omega/kT}-1} = \frac{1}{\tau^2} - \left(\frac{\pi kT}{\hbar}\right)^2 \operatorname{csch}^2\left(\frac{\pi kT\tau}{\hbar}\right) \tag{2.118}$$

and

$$\begin{aligned}\langle\phi(0,t)\phi(0,t+\tau)\rangle &= \frac{\hbar}{\pi c}\left[-\frac{1}{\tau^2}+\frac{1}{\tau^2}-\left(\frac{\pi kT}{\hbar}\right)^2\operatorname{csch}^2\left(\frac{\pi kT\tau}{\hbar}\right)\right] \\ &= -\frac{\hbar}{\pi c}\left(\frac{\pi kT}{\hbar}\right)^2\operatorname{csch}^2\left(\frac{\pi kT\tau}{\hbar}\right).\end{aligned} \tag{2.119}$$

Let us consider also the correlation function $\langle\phi(\mathbf{y},t)\phi(\mathbf{y}+\mathbf{x},t+\tau)\rangle_0$ in the *vacuum* state of our scalar field. In this case $\langle a_{\mathbf{k}}(0)a_{\mathbf{k}'}(0)\rangle = \langle a^\dagger_{\mathbf{k}}(0)a_{\mathbf{k}'}(0)\rangle = 0$ and $\langle a_{\mathbf{k}}(0)a^\dagger_{\mathbf{k}'}(0)\rangle = \delta^3_{\mathbf{k},\mathbf{k}'}$, and from (2.113) we obtain

$$\begin{aligned}\langle\phi(\mathbf{y},t)\phi(\mathbf{y}+\mathbf{x},t+\tau)\rangle_0 &= \sum_{\mathbf{k}}\left(\frac{2\pi\hbar c^2}{\omega_k V}\right)e^{-i\mathbf{k}\cdot\mathbf{x}}e^{i\omega_k\tau} \\ &\to \frac{\hbar c^2}{4\pi^2}\int_0^\infty dk k^2\omega^{-1}e^{i\omega\tau}\int d\Omega_{\mathbf{k}}e^{-i\mathbf{k}\cdot\mathbf{x}},\end{aligned} \tag{2.120}$$

where the last integral is over all solid angles about $\mathbf{k}$:

$$\int d\Omega_{\mathbf{k}}e^{-i\mathbf{k}\cdot\mathbf{x}} = \int_0^{2\pi}d\phi\int_0^\pi d\theta\sin\theta e^{-ikx\cos\theta} = 4\pi\frac{\sin kx}{kx}. \tag{2.121}$$

Thus

$$\begin{aligned}\langle\phi(\mathbf{y},t)\phi(\mathbf{y}+\mathbf{x},t+\tau)\rangle_0 &= \frac{\hbar c^2}{\pi}\int_0^\infty dk k^2\omega^{-1}e^{i\omega\tau}\frac{\sin kx}{kx} \\ &= \frac{\hbar c}{\pi}\frac{1}{x^2-c^2\tau^2}.\end{aligned} \tag{2.122}$$

[10] L. S. Gradshteyn and I. M. Rhyzhik, *Table of Integrals, Series, and Products* (Academic Press, New York, 1980), p. 494, No. 13.



We turn now to an observer undergoing uniform acceleration in the vacuum. Uniform acceleration here is defined with respect to an instantaneous inertial frame in which the observer is at rest. The *proper* acceleration $a$ is the acceleration relative to this instantaneous inertial rest frame, and if $a$ is constant the acceleration is said to be uniform. The acceleration $dv/dt$ in the lab frame may be related to $a$ using standard Lorentz transformations for acceleration:

$$\frac{dv}{dt} = a\left(1-\frac{v^2}{c^2}\right)^{3/2}. \tag{2.123}$$

Simple integrations give $v(t) = at(1+a^2t^2/c^2)^{-1/2}$ and $x(t) = c^2/a[(1+a^2t^2/c^2)^{1/2}-1]$ if we assume $v = x = 0$ when $t = 0$. Using the relation $dt = d\tau(1-v^2/c^2)^{-1/2}$ between lab and proper time intervals, respectively, we have

$$\frac{dt}{d\tau} = \left(1-\frac{a^2t^2}{c^2+a^2t^2}\right)^{-1/2} = \left(1+\frac{a^2t^2}{c^2}\right)^{1/2} \tag{2.124}$$

and

$$t(\tau) = \frac{c}{a}\sinh\frac{a\tau}{c} \tag{2.125}$$

if we define $t(\tau=0)=0$. We can use this result to express $x$ and $v$ in the lab frame in terms of the proper time $\tau$:

$$x(\tau) = \frac{c^2}{a}[\cosh\frac{a\tau}{c}-1], \tag{2.126}$$

$$v(\tau) = c\tanh\frac{a\tau}{c}. \tag{2.127}$$

We recall as an aside the motion of a particle of rest mass $m$ acted upon by a constant force $F$. In this case the linear momentum $p = Ft = mv(1-v^2/c^2)^{-1/2}$ and so $v = (Ft/m)[1+(Ft/mc^2)]^{-1/2}$ and $x = (mc^2/F)([1+(Ft/mc^2)]^{1/2}-1)$, which are the results given previously for $a = F/m$. The world line is a hyperbola in the $x-t$ plane, with asymptote $x = ct$, and consequently this motion is often called *hyperbolic motion*. For $Ft << mc^2$ we have the classical parabolic motion, $x(t) = \frac{1}{2}at^2$.

The vacuum correlation function $\langle\phi(x_1,t_1)\phi(x_2,t_2)\rangle_0$ measured by our uniformly accelerated observer is given by (2.122) with $x = x_2-x_1$ and $\tau = t_2-t_1$, or $x = (c^2/a)[\cosh(a\tau_2/c)-\cosh(a\tau_1/c)]$ and $\tau = (c/a)[\sinh(a\tau_2/c)-\sinh(a\tau_1/c)]$. Since

$$\begin{aligned}x^2-c^2\tau^2 &= \frac{c^4}{a^2}[\cosh\frac{a\tau_2}{c}-\cosh\frac{a\tau_1}{c}]^2-\frac{c^4}{a^2}[\sinh\frac{a\tau_2}{c}-\sinh\frac{a\tau_1}{c}]^2 \\ &= -\frac{4c^4}{a^2}\sinh^2\frac{a(\tau_2-\tau_1)}{2c},\end{aligned} \tag{2.128}$$

it follows from (2.122) that

$$\langle\phi(x_1,t_1)\phi(x_2,t_2)\rangle_0 = -\frac{\hbar a^2}{\pi c^3}\mathrm{csch}^2\frac{a(\tau_2-\tau_1)}{2c}, \tag{2.129}$$

which is equivalent to the *thermal-field* correlation function (2.119) with temperature

$$T = \frac{\hbar a}{2\pi kc}. \tag{2.130}$$

The meaning of this result is that a uniformly accelerated detector in the vacuum responds as it would if it were at rest in a thermal bath at temperature $T = \hbar a/2\pi kc$. In a sense the effect of the acceleration is to "promote" zero-point quantum field fluctuations to the level of thermal fluctuations. It is hardly obvious why this should be so — it took half a century after the birth of the quantum theory of radiation for the thermal effect of uniform acceleration to be discovered.

## 2.11 Thermal Radiation

There are two reasons for reviewing aspects of thermal radiation in this section and the next. First, certain statistical properties of thermal radiation are similar to those of the vacuum field. Second, the quantum theory of thermal radiation provides a clearer picture of some results used in Chapter 1, particularly in connection with the role of the zero-point (vacuum) field in the blackbody problem.

The probability $P_n$ that there are $n$ photons in a field mode of frequency $\omega$ in thermal equilibrium at temperature $T$ — that is, that the mode is excited to the harmonic oscillator level $n$ — is

$$\begin{aligned} P_n &= \frac{e^{-(n+\frac{1}{2})\hbar\omega/kT}}{\sum_{n=0}^{\infty} e^{-(n+\frac{1}{2})\hbar\omega/kT}} = \frac{e^{-n\hbar\omega/kT}}{\sum_{n=0}^{\infty} e^{-n\hbar\omega/kT}} \\ &= e^{-n\hbar\omega/kT}(1-e^{-\hbar\omega/kT})^{-1}\,. \end{aligned} \tag{2.131}$$

The average photon number is thus

$$\overline{n} = \sum_{n=0}^{\infty} nP_n = (e^{\hbar\omega/kT}-1)^{-1}\,, \tag{2.132}$$

and we can use this result to write $P_n$ in terms of $\overline{n}$:

$$P_n = \frac{\overline{n}^n}{(\overline{n}+1)^{n+1}}\,. \tag{2.133}$$



Since there are $(\omega^2/\pi^2c^3)V\,d\omega$ modes of the field in the frequency interval $[\omega,\omega+d\omega]$ in a volume $V$ large compared with $c/\omega$, the spectral energy density is $\rho(\omega) = \hbar\omega\overline{n}(\omega)\omega^2/\pi^2c^3$, which of course is the Planck spectrum, without the zero-point contribution.

The results (2.131)–(2.133) depend only on the frequency of the radiation, not its wave vector $\mathbf{k}$ or polarization $\lambda$. *Thermal radiation as described by the Planck spectrum is isotropic and unpolarized.*

We can use (2.133) to calculate averages of functions of $n$. For instance,

$$\langle n^2\rangle = \sum_{n=0}^{\infty} n^2P_n = 2\overline{n}^2+\overline{n}, \tag{2.134}$$

and so

$$\langle\Delta n^2\rangle \equiv \langle n^2\rangle - \langle n\rangle^2 = 2\overline{n}^2+\overline{n}-\overline{n}^2 = \overline{n}^2+\overline{n}, \tag{2.135}$$

which is a well known consequence of Bose–Einstein statistics.[11]

Since $\rho(\omega) = \hbar\omega^3\overline{n}(\omega)/\pi^2c^3$, we can write (2.135) in the form

$$\begin{aligned} \langle\Delta n(\omega)^2\rangle &= \frac{\pi^2c^3}{\hbar\omega^3}\left[\frac{\pi^2c^3}{\hbar\omega^3}\rho^2(\omega)+\rho(\omega)\right] \\ &= \frac{\pi^2c^3}{\hbar^2\omega^4}\left[\hbar\omega\rho(\omega)+\frac{\pi^2c^3}{\omega^2}\rho^2(\omega)\right]. \end{aligned} \tag{2.136}$$

The variance in the energy of the thermal field is thus

$$\begin{aligned} \langle\Delta E^2\rangle &= \sum_{\mathbf{k}\lambda}\hbar^2\omega_k^2\langle\Delta n(\omega_k)^2\rangle \to \frac{V}{8\pi^3}(2)\int d^3k\hbar^2\omega^2\langle\Delta n(\omega)^2\rangle \\ &= \frac{V}{4\pi^3c^3}(4\pi)\int d\omega\omega^2(\hbar^2\omega^2)\langle\Delta n(\omega)^2\rangle \equiv \int\langle\Delta E_\omega^2\rangle, \end{aligned} \tag{2.137}$$

where

$$\langle\Delta E_\omega^2\rangle = \frac{V}{\pi^2c^3}\hbar^2\omega^4\langle\Delta n(\omega)^2\rangle d\omega = [\hbar\omega\rho(\omega)+\frac{\pi^2c^3}{\omega^2}\rho^2(\omega)]V\,d\omega. \tag{2.138}$$

This is the Einstein fluctuation formula (1.63).

And so Einstein's fluctuation formula can be regarded as a precursor of the result (2.135) of Bose–Einstein statistics. From the discussion in

[11] See, for instance, L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Addison–Wesley, Reading, Mass., 1969), p. 355.

Section 1.7 we can associate $\overline{n}^2$ in (2.135) with wave fluctuations, and $\overline{n}$ with particle fluctuations.

In Section 1.7 we also inferred that the particle fluctuation term could be attributed to the zero-point energy of the field. To appreciate this from the perspective of the quantum theory of the field, note that

$$\begin{aligned}\langle \Delta n^2 \rangle &= \langle n^2 \rangle - \langle n \rangle^2 = \langle a^\dagger a a^\dagger a \rangle - \langle a^\dagger a \rangle^2 \\ &= \langle a^\dagger (a^\dagger a + 1) a \rangle - \langle a^\dagger a \rangle^2 \\ &= \langle a^\dagger a^\dagger a a \rangle + \langle a^\dagger a \rangle - \langle a^\dagger a \rangle^2 \\ &= \langle a^\dagger a^\dagger a a \rangle + \overline{n} - \overline{n}^2, \end{aligned} \tag{2.139}$$

where $a$ is the photon annihilation operator for the field mode under consideration. Now a mode of a thermal field is described by the density matrix

$$\rho = \sum_n P_n |n\rangle\langle n| \,, \tag{2.140}$$

and therefore

$$\langle a^\dagger a^\dagger a a \rangle = \sum_n \langle n | a^\dagger a^\dagger a a | n \rangle. \tag{2.141}$$

But $aa|n\rangle = \sqrt{n}a|n-1\rangle = \sqrt{n(n-1)}|n-2\rangle$, so that

$$\langle a^\dagger a^\dagger a a \rangle = \sum_{n=0}^{\infty} n(n-1) P_n = \sum_{n=0}^{\infty} n(n-1) \frac{\overline{n}^n}{(\overline{n}+1)^{n+1}} = 2\overline{n}^2 \tag{2.142}$$

for a thermal field. Then (2.139) reproduces (2.135). But note that the particle term $\overline{n}$ in this formula arises from the second term in the last line of (2.139), i.e., *from the fact that the commutator* $[a, a^\dagger] = 1$. Note furthermore that this same commutator gives rise to the zero-point energy of a harmonic oscillator such as a field mode, as is clear from equation (2.9). The conclusion is obvious: *the particle term in the Einstein fluctuation formula, or equivalently (2.135), is closely linked to the existence of zero-point energy.*

The "wave" fluctuation term $\overline{n}^2$ in the variance $\langle \Delta n^2 \rangle$ for thermal radiation arises from the factor of 2 in $\langle a^\dagger a^\dagger a a \rangle = 2\overline{n}^2$. This important factor is the origin of Brown–Twiss correlations,[12] also known as *photon bunching*. Suppose we take a spectrally filtered beam of thermal radiation and

[12] R. Hanbury Brown and R. Q. Twiss, *Nature* **127**, 27 (1956); *Proc. Roy. Soc. Lond.* **A242**, 300 (1957). For a discussion of the Brown–Twiss effect see, for instance, Knight and Allen (1983); Loudon (1983); Milonni (1984).



employ a detection scheme in which photons are counted by two-photon absorption rather than ordinary one-photon absorption. That is, a photoelectron is produced by the simultaneous absorption of two photons. As shown in Appendix E, such a detector responds to the *normally ordered* field correlation function $\langle a^\dagger a^\dagger a a \rangle$ if we have a single field mode. The fact that this quantity exceeds $\overline{n}^2$ indicates that the photons have a statistical tendency to arrive in pairs. Such photon bunching of thermal radiation was first measured by Brown and Twiss in the 1950s.

It is worth noting that this "photon bunching" may be understood in purely classical terms, based on the Einstein–Hopf model of a thermal field as a superposition of waves with independent random phases (Section 1.5). Comparing equations (1.67) and (1.68), we note that $\langle \mathbf{E}^4(\mathbf{r},t) \rangle = 2\langle \mathbf{E}^2(\mathbf{r},t) \rangle^2$, or $\overline{I^2} = 2\overline{I}^2$. Thus there are positive intensity correlations or, in photon language, a tendency for photons to arrive in pairs.

It should be emphasized that photon bunching is not a universal property of light. An ideal laser, for example, gives $\langle a^\dagger a^\dagger a a \rangle = \overline{n}^2$, indicating that the photon arrivals are uncorrelated. In other words, an ideal laser has no wave fluctuations: $\langle \Delta n^2 \rangle = \overline{n}$. It is the closest we can get to the idealized, nonfluctuating classical wave of light.

A thermal field, like the vacuum field, is described by Gaussian statistics. Consider for simplicity a single mode of the field, for which the electric field operator is given by equation (2.35). The characteristic function of a single component of this field, which is defined as

$$C[E(\mathbf{r},t), \xi] = \langle e^{i\xi E(\mathbf{r},t)} \rangle, \tag{2.143}$$

gives the probability distribution $P[E(\mathbf{r},t)]$ via a Fourier transform:

$$P[E(\mathbf{r},t)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\xi e^{-i\xi E} C[E, \xi]. \tag{2.144}$$

Using (2.35), we have

$$C[E(\mathbf{r},t), \xi] = \langle e^{i\xi(\alpha a + \alpha^* a^\dagger)} \rangle, \tag{2.145}$$

where $\alpha \equiv i(2\pi\hbar\omega)^{1/2} A_o(\mathbf{r})$. For a thermal field, according to Bloch's theorem for a harmonic oscillator in thermal equilibrium,[13]

$$\langle e^{i\xi(\alpha a + \alpha^* a^\dagger)} \rangle = e^{-\xi^2 |\alpha|^2 (\overline{n} + \frac{1}{2})}, \tag{2.146}$$

[13] F. Bloch, *Z. Phys.* **74**, 295 (1932). See also W. H. Louisell, *Radiation and Noise in Quantum Electronics* (McGraw-Hill, New York, 1964), p. 244.

so that, from (2.144),

$$P[E(\mathbf{r},t)] = \frac{1}{\sqrt{2\pi\mu}} e^{-E^2(\mathbf{r},t)/2\mu}, \tag{2.147}$$

with

$$\mu = 2|\alpha|^2(\overline{n}+\frac{1}{2}) = 4\pi\hbar\omega|A_o(\mathbf{r})|^2(\overline{n}+\frac{1}{2}). \tag{2.148}$$

Comparing (2.147) with the distribution (2.42) for the case of the vacuum field, we see that both the vacuum and thermal fields are distributed according to a Gaussian probability distribution. The vacuum distribution is just the $T \to 0$ limit of the thermal distribution. These results are easy to generalize to the multimode case.

## 2.12 Thermal Equilibrium

We now turn our attention once more to the Einstein–Hopf model of thermal equilibrium between radiation and matter, this time treating both the radiation and the dipole oscillators quantum mechanically (Milonni, 1981a). This will allow us, among other things, to better understand the Einstein–Stern derivation of the Planck spectrum discussed in Section 1.6.

The impulse imparted to a dipole oscillator in the quantum-mechanical version of the Einstein–Hopf model is given by equation (1.48), but now $z(t)$ and $\partial E_z(t)/\partial x$ are quantum-mechanical operators. For the dipole oscillator we introduce lowering and raising operators, $\sigma$ and $\sigma^\dagger$, respectively, as in Section 2.2: $z = i(\hbar/2m\omega_o)^{1/2}(\sigma - \sigma^\dagger)$, $[\sigma,\sigma^\dagger] = 1$. For $\partial E_z/\partial x$ we have, from (2.56),

$$\frac{\partial E_z}{\partial x} = -\sum_{\mathbf{k}\lambda}\left(\frac{2\pi\hbar\omega_k}{V}\right)^{1/2} k_x e_{\mathbf{k}\lambda z}[a_{\mathbf{k}\lambda}(t) + a^\dagger_{\mathbf{k}\lambda}(t)] \tag{2.149}$$

at the position $\mathbf{r} = 0$ of one of the dipole oscillators. Thus

$$\begin{aligned}\Delta &= e\int_0^\tau dt z(t)\frac{\partial E_z(t)}{\partial x} \\ &= -ie\left(\frac{\hbar}{2m\omega_o}\right)^{1/2}\sum_{\mathbf{k}\lambda}\left(\frac{2\pi\hbar\omega_k}{V}\right)^{1/2} k_x e_{\mathbf{k}\lambda z} \\ &\quad\times\int_0^\tau dt[\sigma(t) - \sigma^\dagger(t)][a_{\mathbf{k}\lambda}(t) + a^\dagger_{\mathbf{k}\lambda}(t)].\end{aligned} \tag{2.150}$$



We shall assume that the dipole-field coupling is sufficiently weak over the time interval $[0,\tau]$ that $\sigma(t)$ and $a_{\mathbf{k}\lambda}(t)$ follow approximately their free evolution in (2.150): $\sigma(t) \cong \sigma(0)e^{-i\omega_o t} \equiv \sigma e^{-i\omega_o t}$ and $a_{\mathbf{k}\lambda}(t) \cong a_{\mathbf{k}\lambda}(0)e^{-i\omega_k t} \equiv a_{\mathbf{k}\lambda}e^{-i\omega_k t}$:

$$\begin{aligned}\Delta &\cong -ie\left(\frac{\hbar}{2m\omega_o}\right)^{1/2}\sum_{\mathbf{k}\lambda}\left(\frac{2\pi\hbar\omega_k}{V}\right)^{1/2} k_x e_{\mathbf{k}\lambda z}\left[\sigma a_{\mathbf{k}\lambda}\int_0^\tau dt e^{-i(\omega_k+\omega_o)t}\right. \\ &\quad\left.+\sigma a^\dagger_{\mathbf{k}\lambda}\int_0^\tau dt e^{i(\omega_k-\omega_o)t} - \text{h.c.}\right] \\ &= ie\left(\frac{\hbar}{2m\omega_o}\right)^{1/2}\sum_{\mathbf{k}\lambda}\left(\frac{2\pi\hbar\omega_k}{V}\right)^{1/2} k_x e_{\mathbf{k}\lambda z} \\ &\quad\times\left[\sigma c_{\mathbf{k}\lambda}e^{-i(\omega_k+\omega_o)\tau/2}\frac{\sin\frac{1}{2}(\omega_k+\omega_o)\tau}{\frac{1}{2}(\omega_k+\omega_o)}\right. \\ &\quad\left.-\sigma a^\dagger_{\mathbf{k}\lambda}e^{i(\omega_k-\omega_o)\tau/2}\frac{\sin\frac{1}{2}(\omega_k-\omega_o)\tau}{\frac{1}{2}(\omega_k-\omega_o)} - \text{h.c.}\right].\end{aligned} \tag{2.151}$$

Terms involving $\omega_k + \omega_o$ do not in the end contribute to $\langle\Delta^2\rangle$, just as such "energy nonconserving" terms do not contribute to transition probabilities in standard second-order perturbation theory. Thus

$$\begin{aligned}\Delta &\cong -ie\left(\frac{\hbar}{2m\omega_o}\right)^{1/2}\sum_{\mathbf{k}\lambda}\left(\frac{2\pi\hbar\omega_k}{V}\right)^{1/2} k_x e_{\mathbf{k}\lambda z} \\ &\quad\times[\sigma a^\dagger_{\mathbf{k}\lambda}e^{i(\omega_k-\omega_o)\tau/2} - \text{h.c.}]\frac{\sin\frac{1}{2}(\omega_k-\omega_o)\tau}{\frac{1}{2}(\omega_k-\omega_o)}\end{aligned} \tag{2.152}$$

and

$$\begin{aligned}\langle\Delta^2\rangle &\cong e^2\left(\frac{2\hbar}{m\omega_o}\right)\sum_{\mathbf{k}\lambda}\left(\frac{2\pi\hbar\omega_k}{V}\right)k_x^2 e^2_{\mathbf{k}\lambda z}\frac{\sin^2\frac{1}{2}(\omega_k-\omega_o)\tau}{(\omega_k-\omega_o)^2} \\ &\quad\times\left[\langle\sigma\sigma^\dagger\rangle\langle a^\dagger_{\mathbf{k}\lambda}a_{\mathbf{k}\lambda}\rangle + \langle\sigma^\dagger\sigma\rangle\langle a_{\mathbf{k}\lambda}a^\dagger_{\mathbf{k}\lambda}\rangle\right].\end{aligned} \tag{2.153}$$

Let us write $\langle a^\dagger_{\mathbf{k}\lambda}a_{\mathbf{k}\lambda}\rangle = \overline{n}(\omega)$ and $\langle a_{\mathbf{k}\lambda}a^\dagger_{\mathbf{k}\lambda}\rangle = \langle a^\dagger_{\mathbf{k}\lambda}a_{\mathbf{k}\lambda}\rangle + 1 = \overline{n}(\omega)+1$, and proceed to the mode continuum limit $V \to \infty$ in (2.153):

$$\langle\Delta^2\rangle \cong e^2\left(\frac{2\hbar}{m\omega_o}\right)\left(\frac{V}{8\pi^3}\right)\int d^3k\left(\frac{2\pi\hbar\omega}{V}\right)k_x^2\sum_\lambda e^2_{\mathbf{k}\lambda z}\frac{\sin^2\frac{1}{2}(\omega-\omega_o)\tau}{(\omega-\omega_o)^2}$$

$$\times \left[\langle\sigma\sigma^\dagger\rangle\overline{n}(\omega) + \langle\sigma^\dagger\sigma\rangle(\overline{n}(\omega)+1)\right]$$
$$= \frac{e^2\hbar^2}{2\pi^2 m\omega_o}\int d^3k\omega k_x^2\left(1-\frac{k_z^2}{k^2}\right)\frac{\sin^2\frac{1}{2}(\omega-\omega_o)\tau}{(\omega-\omega_o)^2}$$
$$\times \left[\langle\sigma\sigma^\dagger\rangle\overline{n}(\omega) + \langle\sigma^\dagger\sigma\rangle(\overline{n}(\omega)+1)\right]. \qquad (2.154)$$

In writing this expression we have used the identity $\hat{\mathbf{z}} = (\hat{\mathbf{k}}\cdot\hat{\mathbf{z}})\hat{\mathbf{k}} + \sum_\lambda e_{\mathbf{k}\lambda z}\mathbf{e}_{\mathbf{k}\lambda}$ for any unit vector $\hat{\mathbf{z}}$, and therefore $\sum_\lambda e_{\mathbf{k}\lambda z}^2 = 1-(\hat{\mathbf{k}}\cdot\hat{\mathbf{z}})^2 = 1-k_z^2/k^2$. Now

$$\int d^3k k_x^2\left(1-\frac{k_z^2}{k^2}\right) = \int dk k^2\int d\Omega_{\mathbf{k}} k_x^2\left(\frac{k_x^2+k_y^2}{k^2}\right)$$
$$= \int dk\int d\Omega_{\mathbf{k}} k_x^2(k_x^2+k_y^2)$$
$$= \int dk k^4\int_0^{2\pi}d\phi\int_0^\pi d\theta\sin\theta(\sin\theta\cos\phi)^2\sin^2\theta$$
$$= \frac{16\pi}{15}\int dk k^4 = \frac{16\pi}{15c^5}\int d\omega\omega^4 \qquad (2.155)$$

and so

$$\langle\Delta^2\rangle = \frac{8e^2\hbar^2}{15\pi m\omega_o c^5}\int_0^\infty d\omega\omega^5\frac{\sin^2\frac{1}{2}(\omega-\omega_o)\tau}{(\omega-\omega_o)^2}$$
$$\times\left[\langle\sigma\sigma^\dagger\rangle\overline{n}(\omega)+\langle\sigma^\dagger\sigma\rangle(\overline{n}(\omega)+1)\right]$$
$$\cong \frac{8e^2\hbar^2\omega_o^4}{15\pi mc^5}\left[\langle\sigma\sigma^\dagger\rangle\overline{n}(\omega_o)+\langle\sigma^\dagger\sigma\rangle(\overline{n}(\omega_o)+1)\right]$$
$$\times\int_0^\infty d\omega\frac{\sin^2\frac{1}{2}(\omega-\omega_o)\tau}{(\omega-\omega_o)^2}$$
$$\cong \frac{4e^2\hbar^2\omega_o^4}{15mc^5}\left[\langle\sigma\sigma^\dagger\rangle\overline{n}(\omega_o)+\langle\sigma^\dagger\sigma\rangle(\overline{n}(\omega_o)+1)\right]\tau. \qquad (2.156)$$

As in the classical Einstein–Hopf model the condition for thermal equilibrium is $\tau^{-1}\langle\Delta^2\rangle = 2RkT$ [Equation (1.47)]. From (2.156) and (1.43) this condition is

$$\rho(\omega_o) - \frac{\omega_o}{3}\frac{d\rho}{d\omega_o} = \frac{\hbar^2\omega_o^4}{6\pi^2c^3kT}\left[\langle\sigma\sigma^\dagger\rangle\overline{n}(\omega_o)+\langle\sigma^\dagger\sigma\rangle(\overline{n}(\omega_o)+1)\right]. \qquad (2.157)$$



Now the dipole is just another harmonic oscillator, and so in thermal equilibrium $\langle\sigma\sigma^\dagger\rangle = \langle\sigma^\dagger\sigma\rangle + 1 = \overline{n}(\omega_o)+1$, whence

$$\langle\sigma\sigma^\dagger\rangle\overline{n}(\omega_o)+\langle\sigma^\dagger\sigma\rangle[\overline{n}(\omega_o)+1] = 2\overline{n}(\omega_o)[\overline{n}(\omega_o)+1]$$
$$= 2\left(\frac{\pi^2c^3}{\hbar\omega_o^3}\right)^2[\rho^2(\omega_o)+\frac{\hbar\omega_o^3}{\pi^2c^3}\rho(\omega_o)], \qquad (2.158)$$

where we have again employed the relation $\overline{n}(\omega_o) = (\pi^2c^3/\hbar\omega_o^3)\rho(\omega_o)$. Then equation (2.157) yields exactly the Einstein–Stern equation (1.55), whose solution is the Planck spectrum.

It is hardly surprising that the quantum theory of the Einstein–Hopf model produces the Planck spectrum for the spectral energy density of radiation at thermal equilibrium. What is of interest is to see just what about the quantum theory leads to the Planck spectrum rather than the Rayleigh–Jeans spectrum of the classical Einstein–Hopf model. To this end we use the identities $\sigma\sigma^\dagger = \sigma^\dagger\sigma + 1$ and $a_{\mathbf{k}\lambda}a_{\mathbf{k}\lambda}^\dagger = a_{\mathbf{k}\lambda}^\dagger a_{\mathbf{k}\lambda} + 1$ to write

$$\langle\sigma\sigma^\dagger\rangle\langle a_{\mathbf{k}\lambda}^\dagger a_{\mathbf{k}\lambda}\rangle + \langle\sigma^\dagger\sigma\rangle\langle a_{\mathbf{k}\lambda}a_{\mathbf{k}\lambda}^\dagger\rangle = 2\left[\langle\sigma^\dagger\sigma\rangle\langle a_{\mathbf{k}\lambda}^\dagger a_{\mathbf{k}\lambda}\rangle + \frac{1}{2}\langle\sigma^\dagger\sigma\rangle + \frac{1}{2}\langle a_{\mathbf{k}\lambda}^\dagger a_{\mathbf{k}\lambda}\rangle\right]. \qquad (2.159)$$

Without the term $\frac{1}{2}\langle\sigma^\dagger\sigma\rangle + \frac{1}{2}\langle a_{\mathbf{k}\lambda}^\dagger a_{\mathbf{k}\lambda}\rangle$ in this expression we are led to the Rayleigh–Jeans spectrum. In other words, the Planck spectrum is a consequence of the quantum-mechanical commutation rules $[\sigma,\sigma^\dagger] = [a_{\mathbf{k}\lambda}, a_{\mathbf{k}\lambda}^\dagger] = 1$.

For a more physical interpretation of the role of quantum mechanics, let us note that in the final expression for $\tau^{-1}\langle\Delta^2\rangle$ only the field modes at $\omega_k = \omega_o$ contribute. These modes impart a mean-square momentum transfer proportional to

$$(\hbar\omega_o)^2\left[\langle\sigma^\dagger\sigma\rangle\overline{n}(\omega_o)+\frac{1}{2}\langle\sigma^\dagger\sigma\rangle+\frac{1}{2}\overline{n}(\omega_o)\right] = \langle H_{\rm osc}\rangle\langle H_{\rm F}\rangle + H_{\rm F}^{\rm zp}\langle H_{\rm osc}\rangle + H_{\rm osc}^{\rm zp}\langle H_{\rm F}\rangle, \qquad (2.160)$$

where $H_{\rm osc} = \hbar\omega_o\langle\sigma^\dagger\sigma\rangle$ and $H_{\rm F} = \hbar\omega_o a^\dagger a$ are the Hamiltonian operators for the dipole oscillator and a resonant field mode, respectively, excluding zero-point energies, and $H_{\rm osc}^{\rm zp}$ and $H_{\rm F}^{\rm zp}$ are the corresponding zero-point energies ($= \frac{1}{2}\hbar\omega_o$). Were it not for the zero-point energies in (2.160), we

would obtain the classical Rayleigh- Jeans spectrum instead of the Planck spectrum. These terms give rise to the "particle" term proportional to $\rho(\omega_o)$ in (2.158).

All this is consistent with the classical discussion in Section 1.6. Since

$$H_{\rm F}^{\rm zp}\langle H_{\rm osc}\rangle + H_{\rm osc}^{\rm zp}\langle H_{\rm F}\rangle = 2H_{\rm osc}^{\rm zp}\langle H_{\rm F}\rangle = \hbar\omega_o\langle H_{\rm F}\rangle, \qquad (2.161)$$

a classical theory of the Einstein–Hopf model that includes a zero-point energy $\hbar\omega_o$ for a material oscillator, but not for any field oscillator, leads to the same equation (2.157) of quantum theory, and therefore gives the Planck spectrum. This was the approach of Einstein and Stern.

Alternatively, we can include in the classical theory a zero-point energy $\frac{1}{2}\hbar\omega_o$ for both the material oscillator and a field mode of frequency $\omega_o$, and this too leads to the Planck spectrum, as discussed in Section 1.6. This approach is closer to the (quantum-mechanical) truth. But in such a classical approach we must follow the ad hoc procedure of dropping a contribution $\frac{1}{2}\hbar\omega_o\rho_o(\omega_o)$, which arises from the product $H_{\rm osc}^{\rm zp}H_{\rm F}^{\rm zp}$. There was really no justification of this *Ansatz* in Section 1.6 other than the fact that it gave the Planck spectrum.

In the quantum theory just presented, the terms $H_{\rm F}^{\rm zp}\langle H_{\rm osc}\rangle$ and $H_{\rm osc}^{\rm zp}\times\langle H_{\rm F}\rangle$ leading to the Planck spectrum arise "automatically" from the zero-point energies of the dipole and field oscillators or, more formally, from the commutation properties of the dipole and field operators. But there is no term $H_{\rm osc}^{\rm zp}H_{\rm F}^{\rm zp}$ that had to be dropped ad hoc in the classical approach to the Planck spectrum presented in Section 1.6. In other words, *the quantum theory of the Einstein–Hopf model apparently does not allow for any effect of the interaction between a ground-state dipole oscillator and the vacuum field.*

We must be careful here about what we mean by the "effect" of the vacuum field on a ground- state dipole oscillator. The dipole coordinate obeys the Heisenberg equation of motion (2.86), and we have shown that the vacuum field is necessary for the preservation of the canonical commutation relations for the dipole coordinate and momentum operators, regardless of the state of the dipole. In this sense the vacuum field certainly has a formal "effect." Physically, however, a dipole oscillator in its ground state shows no obvious effect of its interaction with the vacuum field: a ground-state oscillator in the vacuum remains forever in its ground state. Whereas an excited dipole oscillator can undergo spontaneous emission attributable in part to the vacuum field, there is no such thing as "spontaneous absorption" by a ground-state oscillator in vacuum. We shall see that in the ground state of an atom spontaneous absorption is precluded by an exact cancellation of vacuum field fluctuations by fluctuations in the atom.



## 2.13 Summary

In the quantum theory of the electromagnetic field, classical wave amplitudes $\alpha, \alpha^*$ are replaced by operators $a, a^\dagger$ satisfying $[a, a^\dagger] = 1$. The quantity $|\alpha|^2$ appearing in the classical expression for the energy of a field mode [cf. equation (2.29)] is replaced in quantum theory by the photon number operator $a^\dagger a$. The fact that $[a, a^\dagger a] \neq 0$ implies that quantum theory does not allow states of the radiation field for which the photon number and a field amplitude can be precisely defined, i.e., we cannot have simultaneous eigenstates of $a^\dagger a$ and $a$. The reconciliation of wave and particle attributes of the field is accomplished via the association of a probability *amplitude* with a classical mode pattern, as discussed in Section 2.4. The calculation of field modes is an entirely classical problem, while the quantum properties of the field are carried by the mode "amplitudes" $a$ and $a^\dagger$ associated with these classical modes.

The zero-point energy of the field arises formally from the noncommutativity of $a$ and $a^\dagger$. This is true for *any* harmonic oscillator: the zero-point energy $\frac{1}{2}\hbar\omega$ appears when we write the Hamiltonian $H = p^2/2m + \frac{1}{2}m\omega^2q^2 = \frac{1}{2}\hbar\omega(aa^\dagger + a^\dagger a)$ as $\hbar\omega(a^\dagger a + \frac{1}{2})$ [Equation (2.9)].

This zero-point energy can be dropped from the Hamiltonian by redefining the zero of energy, or by arguing that it is a c-number and therefore has no effect on Heisenberg equations of motion. However, when we do this and solve the Heisenberg equation for a field operator, we must include the vacuum field, which is the homogeneous part of the solution for the field operator. In fact we showed in Section 2.6 that the vacuum field is essential for the preservation of commutators and the formal consistency of the theory. When we calculate the field energy we obtain not only a contribution from any sources which may be present, but also a contribution from the vacuum field. The latter is of course the zero-point field energy. In other words, the zero-point field energy "reappears" even though we may have deleted it from the Hamiltonian.

As we saw in the first chapter, the concept of zero-point energy arose before the development of the quantum formalism. However, in quantum theory zero-point energy rests upon a much firmer foundation than was possible classically. This is illustrated by a comparison of the Einstein–Stern theory of blackbody radiation with the quantum theory presented in Section 2.12.

Observable phenomena like the Casimir effect strongly suggest that the vacuum electromagnetic field and its zero-point energy are real physical entities and not mere artifices of the quantum formalism. In the following

chapter we shall turn to other things that similarly suggest the physical reality of the fluctuating vacuum electromagnetic field.

Finally, the Maxwell equations (2.18)–(2.21) are satisfied by the electric and magnetic field *operators* in the quantum theory of the field. Maxwell was lucky: his equations turned out to be Lorentz–invariant and gauge–invariant, and to retain the same form in quantum theory. But whereas in classical physics one makes the "natural" assumption that $\mathbf{E} = \mathbf{B} = 0$ in the absence of any sources, this cannot be done in quantum theory. Such an assumption is not only inconsistent with quantum theory; it would also appear to contradict experimental facts such as the Casimir force, the Lamb shift, and other effects to which we turn in the following chapter.

## 2.14 Bibliography

There are many excellent books on the quantum theory of radiation and quantum electrodynamics. The books by Power and Loudon are particularly readable and succinct introductions to the nonrelativistic theory. Power's book was one of the first to address the zero-point field in any sort of detail. Heitler's book is an old standard and still contains much useful material, while the treatise by Itzykson and Zuber offers a more modern and advanced perspective. Boyer's review of the role of zero-point field energy in long-range forces is also recommended.

Bohr, N. and L. Rosenfeld, "Field and Charge Measurements in Quantum Electrodynamics," *Phys. Rev.* **78**, 794 (1950).

Born, M., W. Heisenberg, and P. Jordan, "Zur Quantenmechanik II," *Z. Phys.* **35**, 557 (1926).

Boyer, T. H., "Derivation of the Blackbody Radiation Spectrum Without Quantum Assumptions," *Phys. Rev.* **182**, 1374 (1969).

Boyer, T. H., "Quantum Zero-Point Energy and Long-Range Forces," *Ann. Phys.* (New York) **56**, 474 (1970).

Casimir, H. B. G., "On the Attraction between Two Perfectly Conducting Plates," *Proc. K. Ned. Akad. Wet.* **51**, 793 (1948).

Cresser, J. D., "Electric Field Commutation Relation in the Presence of a Dipole Atom," *Phys. Rev.* **A29**, 1984 (1984).

Davies, P. C. W., "Scalar Particle Production in Schwarzschild and Rindler Metrics," *J. Phys.* **A8**, 609 (1975).

Dirac, P. A. M., "The Quantum Theory of the Emission and Absorption of Radiation," *Proc. Roy. Soc. Lond.* **A114**, 243 (1927).

Heitler, W., *The Quantum Theory of Radiation*, 3rd ed. (Oxford University Press, London, 1966).

Itzykson, C. and J.-B. Zuber, *Quantum Field Theory* (McGraw–Hill, New York, 1980).

Knight, P. L. and L. Allen, *Concepts of Quantum Optics* (Pergamon Press, Oxford, 1983).

Loudon, R., *The Quantum Theory of Light*, 2nd ed. (Oxford University Press, London, 1983).

Milonni, P. W., "Quantum Mechanics of the Einstein–Hopf Model," *Am. J. Phys.* **49**, 177 (1981a).

Milonni, P. W., "Radiation Reaction and the Nonrelativistic Theory of the Electron," *Phys. Lett.* **82A**, 225 (1981b).

Milonni, P. W., "Casimir Forces Without the Vacuum Radiation Field," *Phys. Rev.* **A25**, 1315 (1982).

Milonni, P. W., "Wave-Particle Duality of Light: A Current Perspective," in *The Wave-Particle Dualism*, ed. S. Diner, D. Fargue, G. Lochak, and F. Selleri (Reidel, Dordrecht, 1984).

Milonni, P. W. and M.-L. Shih, "Casimir Forces," *Contemp. Phys.* **33**, 313 (1993).

Power, E. A., *Introductory Quantum Electrodynamics* (Longmans, London, 1964).

Senitzky, I. R., "Dissipation in Quantum Mechanics: The Harmonic Oscillator," *Phys. Rev.* **119**, 670 (1960).

Unruh, W. G., "Notes on Black-Hole Evaporation," *Phys. Rev.* **D14**, 870 (1976).