CHAPTER 2

THE QUANTUM THEORY OF RADIATION

2–1. CLASSICAL RADIATION FIELD

Transversality condition. In order to study radiation phenomena in the quantum domain we shall first discuss in detail the properties of the vector potential satisfying

$$\nabla \cdot \mathbf{A} = 0 \tag{2.1}$$

within the framework of classical electrodynamics. Equation (2.1) is known as the *transversality* condition; it should not be confused with the Lorentz condition (1.81). The electric or magnetic fields derivable from vector potentials satisfying (2.1) are called *transverse fields* or *radiation fields*. Often the term "transverse field" or "radiation field" is used to refer to a vector potential itself satisfying (2.1).

The transversality condition (2.1) is of interest under a variety of circumstances. First, suppose $j_\mu = 0$. We can then consider a gauge transformation that eliminates the fourth component of A_μ and makes $\mathbf{A}$ obey (2.1). Consider

$$\begin{aligned} \mathbf{A} &\rightarrow \mathbf{A}' = \mathbf{A} + \nabla\Lambda, \\ A_0 &\rightarrow A_0' = A_0 + \frac{1}{c}\frac{\partial \Lambda}{\partial t}, \end{aligned} \tag{2.2}$$

such that

$$-\frac{1}{c}\frac{\partial \Lambda}{\partial t} = A_0. \tag{2.3}$$

Since the fourth component of A_μ has been eliminated in the new gauge, the Lorentz condition (1.81) reduces to the transversality condition (2.1).

Let us consider a situation in which $j_\mu \neq 0$, as in the case of mutually interacting electrons. We may first decompose $\mathbf{A}$ so that

$$\begin{aligned} \mathbf{A} = \mathbf{A}_\perp + \mathbf{A}_{||}, \qquad \nabla \cdot \mathbf{A}_\perp = 0, \\ \nabla \times \mathbf{A}_{||} = 0. \end{aligned} \tag{2.4}$$

This can always be done.‡ Here $\mathbf{A}_\perp$ and $\mathbf{A}_{||}$ are called respectively the *transverse* and the *longitudinal component* of $\mathbf{A}$. In 1930 E. Fermi was able to show that $\mathbf{A}_{||}$ and A_0 *together* give rise to the instantaneous static Coulomb interactions between

‡See, for example, Morse and Feshbach (1953), pp. 52–54.

the charged particles, whereas $\mathbf{A}_\perp$ accounts for the electromagnetic radiation of moving charged particles. The total Hamiltonian of the charged particles (treated nonrelativistically) and the electromagnetic fields generated by them can be written (apart from the magnetic moment interaction) as:

$$H = \sum_j \frac{1}{2m_j}\left[\mathbf{p}^{(j)} - e_j\frac{\mathbf{A}_\perp(\mathbf{x}^{(j)})}{c}\right]^2 + \sum_{i>j}\frac{e_i e_j}{4\pi\,|\mathbf{x}^{(i)} - \mathbf{x}^{(j)}|} + H_{\text{rad}}, \tag{2.5}$$

where H_{rad} (which we shall discuss later in detail) is the free field Hamiltonian of $\mathbf{A}_\perp$ only. Equation (2.5) is derived in Appendix A. Note that nowhere in (2.5) do A_0 and $\mathbf{A}_{||}$ appear explicitly.

Fermi's formalism based on (2.5) is called the *radiation* (or *Coulomb*) *gauge method.* Since the decomposition (2.5) is not relativistically covariant, nor is the transversality condition itself, the whole formalism appears noncovariant; each time we perform a Lorentz transformation, we must simultaneously make a gauge transformation to obtain a new set of $\mathbf{A}$ and A_0. Yet it is possible to develop manifestly covariant calculational techniques starting with the relativistic analog of the Hamiltonian (2.5), as will be shown in Chapter 4 when we discuss Møller (electron-electron) scattering. It is also possible to construct a formalism which preserves relativistic covariance at every stage.

In any case, it is worth studying a theory of transverse electromagnetic fields before we learn about more sophisticated formalisms. When the theory is quantized, it provides simple and physically transparent descriptions of a variety of processes in which real photons are emitted, absorbed, or scattered. The three basic equations we work with for the free-field case are

$$\mathbf{B} = \nabla \times \mathbf{A}, \qquad \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \tag{2.6}$$

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = 0, \tag{2.7}$$

where $\mathbf{A}$ satisfies the transversality condition (2.1).

Fourier decomposition and radiation oscillators. At a given instant, say $t = 0$, we expand $\mathbf{A}$ in Fourier series. We assume the periodic boundary conditions for $\mathbf{A}$ enclosed in a box taken to be a cube of side $L = (V)^{1/3}$. Remembering the reality of $\mathbf{A}$, we have

$$\mathbf{A}(\mathbf{x}, t)\Big|_{t=0} = \frac{1}{\sqrt{V}}\sum_{\mathbf{k}}\sum_{\alpha=1,2}(c_{\mathbf{k},\alpha}(0)\mathbf{u}_{\mathbf{k},\alpha}(\mathbf{x}) + c^*_{\mathbf{k},\alpha}(0)\mathbf{u}^*_{\mathbf{k},\alpha}(\mathbf{x})), \tag{2.8}$$

where

$$\mathbf{u}_{\mathbf{k},\alpha}(\mathbf{x}) = \boldsymbol{\epsilon}^{(\alpha)}e^{i\mathbf{k}\cdot\mathbf{x}} \tag{2.9}$$

and $\boldsymbol{\epsilon}^{(\alpha)}$, called a (linear) *polarization vector*, is a real unit vector whose direction depends on the propagation direction $\mathbf{k}$ (even though we will not write it in this text as $\boldsymbol{\epsilon}^{(\alpha)}(\mathbf{k})$ as is done in many books). Given $\mathbf{k}$, we choose $\boldsymbol{\epsilon}^{(1)}$ and $\boldsymbol{\epsilon}^{(2)}$ in such a way that $(\boldsymbol{\epsilon}^{(1)}, \boldsymbol{\epsilon}^{(2)}, \mathbf{k}/|\mathbf{k}|)$ form a right-handed set of mutually orthogonal unit vectors. Note that $\boldsymbol{\epsilon}^{(1)}$ and $\boldsymbol{\epsilon}^{(2)}$ are, in general, not along the x- and the y- axes since

$\mathbf{k}$ is, in general, not along the z-axis. Since $\boldsymbol{\epsilon}^{(\alpha)}$ is perpendicular to $\mathbf{k}$, the transversality condition is guaranteed. The Fourier component $\mathbf{u}_{\mathbf{k},\alpha}$ satisfies

$$\frac{1}{V}\int d^3x\, \mathbf{u}_{\mathbf{k},\alpha}\cdot\mathbf{u}^*_{\mathbf{k}',\alpha'} = \delta_{\mathbf{k}\mathbf{k}'}\delta_{\alpha\alpha'},$$
$$\frac{1}{V}\int d^3x \begin{Bmatrix}\mathbf{u}_{\mathbf{k},\alpha}\cdot\mathbf{u}_{\mathbf{k}',\alpha'}\\ \mathbf{u}^*_{\mathbf{k},\alpha}\cdot\mathbf{u}^*_{\mathbf{k}',\alpha'}\end{Bmatrix} = \delta_{\mathbf{k},-\mathbf{k}'}\delta_{\alpha\alpha'}, \tag{2.10}$$

where

$$k_x, k_y, k_z = 2n\pi/L, \qquad n = \pm 1, \pm 2, \ldots \tag{2.11}$$

because of the periodic boundary conditions.

To obtain $\mathbf{A}(\mathbf{x}, t)$ for $t \neq 0$, we simply replace $c_{\mathbf{k},\alpha}(0)$ and $c^*_{\mathbf{k},\alpha}(0)$ by

$$c_{\mathbf{k},\alpha}(t) = c_{\mathbf{k},\alpha}(0)e^{-i\omega t}, \qquad c^*_{\mathbf{k},\alpha}(t) = c^*_{\mathbf{k},\alpha}(0)e^{i\omega t}, \tag{2.12}$$

where

$$\omega = |\mathbf{k}|c. \tag{2.13}$$

With this replacement both the wave equation (2.7) and the reality condition on $\mathbf{A}$ are satisfied. So

$$\mathbf{A}(\mathbf{x}, t) = \frac{1}{\sqrt{V}}\sum_{\mathbf{k}}\sum_{\alpha}(c_{\mathbf{k},\alpha}(t)\boldsymbol{\epsilon}^{(\alpha)}e^{i\mathbf{k}\cdot\mathbf{x}} + c^*_{\mathbf{k},\alpha}(t)\boldsymbol{\epsilon}^{(\alpha)}e^{-i\mathbf{k}\cdot\mathbf{x}})$$
$$= \frac{1}{\sqrt{V}}\sum_{\mathbf{k}}\sum_{\alpha}(c_{\mathbf{k},\alpha}(0)\boldsymbol{\epsilon}^{(\alpha)}e^{ik\cdot x} + c^*_{\mathbf{k},\alpha}(0)\boldsymbol{\epsilon}^{(\alpha)}e^{-ik\cdot x}) \tag{2.14}$$

with

$$k\cdot x = \mathbf{k}\cdot\mathbf{x} - \omega t = \mathbf{k}\cdot\mathbf{x} - |\mathbf{k}|ct. \tag{2.15}$$

The Hamiltonian of the field is

$$H = \tfrac{1}{2}\int (|\mathbf{B}|^2 + |\mathbf{E}|^2)d^3x$$
$$= \tfrac{1}{2}\int [|\nabla\times\mathbf{A}|^2 + |(1/c)(\partial\mathbf{A}/\partial t)|^2]d^3x. \tag{2.16}$$

A typical term we must evaluate for the $|\mathbf{B}|^2$ integration is

$$\int(\nabla\times\mathbf{u}_{\mathbf{k},\alpha})\cdot(\nabla\times\mathbf{u}^*_{\mathbf{k}',\alpha'})d^3x = \int\nabla\cdot[\mathbf{u}_{\mathbf{k},\alpha}\times(\nabla\times\mathbf{u}^*_{\mathbf{k}',\alpha'})]d^3x$$
$$+\int\mathbf{u}_{\mathbf{k},\alpha}\cdot[\nabla\times(\nabla\times\mathbf{u}^*_{\mathbf{k}',\alpha'})]d^3x$$
$$= -\int\mathbf{u}_{\mathbf{k},\alpha}\nabla^2\mathbf{u}^*_{\mathbf{k}',\alpha'}d^3x = \left(\frac{\omega}{c}\right)^2 V\delta_{\mathbf{k}\mathbf{k}'}\delta_{\alpha\alpha'}, \tag{2.17}$$

where we have used the periodic boundary conditions and the identity $\nabla\times(\nabla\times\ ) = \nabla(\nabla\cdot\ ) - \nabla^2$. Similarly, for the $\frac{1}{2}|\mathbf{E}|^2$ integration it is useful to evaluate first

$$\int\left[\frac{1}{c}\frac{\partial}{\partial t}(c_{\mathbf{k},\alpha}\mathbf{u}_{\mathbf{k},\alpha})\right]\cdot\left[\frac{1}{c}\frac{\partial}{\partial t}(c^*_{\mathbf{k}',\alpha'}\mathbf{u}^*_{\mathbf{k}',\alpha'})\right]d^3x = \left(\frac{\omega}{c}\right)^2 V\delta_{\mathbf{k}\mathbf{k}'}\delta_{\alpha\alpha'}c_{\mathbf{k},\alpha}c^*_{\mathbf{k}',\alpha'} \tag{2.18}$$

Using these relations, we obtain

$$H = \sum_{\mathbf{k}} \sum_{\alpha} 2(\omega/c)^2 c^*_{\mathbf{k},\alpha} c_{\mathbf{k},\alpha}, \tag{2.19}$$

where $c_{\mathbf{k},\alpha}$ is a time-dependent Fourier coefficient satisfying

$$\ddot{c}_{\mathbf{k},\alpha} = -\omega^2 c_{\mathbf{k},\alpha} \tag{2.20}$$

(cf. Eq. 2.12). This reminds us of an expression for the energy in a collection of independent and uncoupled harmonic oscillators. To make the analogy more vivid, we define

$$Q_{\mathbf{k},\alpha} = \frac{1}{c}(c_{\mathbf{k},\alpha} + c^*_{\mathbf{k},\alpha}), \qquad P_{\mathbf{k},\alpha} = -\frac{i\omega}{c}(c_{\mathbf{k},\alpha} - c^*_{\mathbf{k},\alpha}). \tag{2.21}$$

Then

$$\begin{aligned} H &= \sum_{\mathbf{k}} \sum_{\alpha} 2\left(\frac{\omega}{c}\right)^2 \left[\frac{c(\omega Q_{\mathbf{k},\alpha} - iP_{\mathbf{k},\alpha})}{2\omega}\right]\left[\frac{c(\omega Q_{\mathbf{k},\alpha} + iP_{\mathbf{k},\alpha})}{2\omega}\right] \\ &= \sum_{\mathbf{k}} \sum_{\alpha} \tfrac{1}{2}(P^2_{\mathbf{k},\alpha} + \omega^2 Q^2_{\mathbf{k},\alpha}). \end{aligned} \tag{2.22}$$

The $P_{\mathbf{k},\alpha}$ and $Q_{\mathbf{k},\alpha}$ are now seen to be canonical variables:

$$\frac{\partial H}{\partial Q_{\mathbf{k},\alpha}} = -\dot{P}_{\mathbf{k},\alpha}, \qquad \frac{\partial H}{\partial P_{\mathbf{k},\alpha}} = +\dot{Q}_{\mathbf{k},\alpha}. \tag{2.23}$$

Thus the radiation field can be regarded as a collection of independent harmonic oscillators each of which is characterized by $\mathbf{k}$, α and whose dynamical variables are orthogonal linear combinations of the Fourier coefficients.

2–2. CREATION, ANNIHILATION, AND NUMBER OPERATORS

Quantization of radiation oscillators. At the end of the nineteenth century it was recognized that the space-time development of the radiation field satisfying the wave equation (2.7) resembles the dynamical behavior of a collection of harmonic oscillators. By assigning an average energy kT to each radiation oscillator, Lord Rayleigh and J. H. Jeans wrote an expressioon for the energy distribution of the radiation field as a function of ω in an ideal situation where the radiation field is enclosed by perfectly absorbing walls. The expression they obtained was in satisfactory agreement with observation for low values of ω at sufficiently high temperatures, but in marked disagreement for high values of ω. This difficulty led M. Planck to take one of the most revolutionary steps ever taken in the history of science. He proposed that:

> The energy of each radiation oscillator is not an arbitrary quantity but must be an integral multiple of $\hbar\omega$, where $\hbar$ is a *new* fundamental constant in nature.

This he did in 1901. Four years later, in order to explain the photoelectric effect, A. Einstein proposed that an electromagnetic wave of wavelength $\lambda = 2\pi c/\omega$ be regarded as a collection of massless particles each of which has energy $\hbar\omega$.

We can now do better than did Planck and Einstein, but only because we know nonrelativistic quantum mechanics. Indeed, no sooner was nonrelativistic quantum mechanics fully developed than P. A. M. Dirac proposed that the canonical dynamical variables of a radiation oscillator be treated as noncommutable operators, just as x and p of a one-dimensional harmonic oscillator are treated in nonrelativistic quantum mechanics.

We postulate that P and Q of the radiation oscillators are no longer mere numbers but are *operators* satisfying

$$[Q_{\mathbf{k},\alpha}, P_{\mathbf{k}',\alpha'}] = i\hbar\delta_{\mathbf{k}\mathbf{k}'}\delta_{\alpha\alpha'}, \tag{2.24a}$$

$$\begin{aligned}[Q_{\mathbf{k},\alpha}, Q_{\mathbf{k}',\alpha'}] &= 0,\\ [P_{\mathbf{k},\alpha}, P_{\mathbf{k}',\alpha'}] &= 0.\end{aligned} \tag{2.24b}$$

We next consider linear combinations of $P_{\mathbf{k},\alpha}$ and $Q_{\mathbf{k},\alpha}$ given by

$$\begin{aligned}a_{\mathbf{k},\alpha} &= (1/\sqrt{2\hbar\omega})(\omega Q_{\mathbf{k},\alpha} + iP_{\mathbf{k},\alpha}),\\ a^\dagger_{\mathbf{k},\alpha} &= (1/\sqrt{2\hbar\omega})(\omega Q_{\mathbf{k},\alpha} - iP_{\mathbf{k},\alpha}).\end{aligned} \tag{2.25}$$

Thus $a_{\mathbf{k},\alpha}$ and $a^\dagger_{\mathbf{k},\alpha}$ are seen to be the operator analogs of the Fourier coefficients $c_{\mathbf{k},\alpha}$ and $c^\dagger_{\mathbf{k},\alpha}$ when we insert a factor to make $a_{\mathbf{k},\alpha}$ and $a^\dagger_{\mathbf{k},\alpha}$ dimensionless:

$$c_{\mathbf{k},\alpha} \to c\sqrt{\hbar/2\omega}\, a_{\mathbf{k},\alpha}.$$

They satisfy the communication relations

$$\begin{aligned}[a_{\mathbf{k},\alpha}, a^\dagger_{\mathbf{k}',\alpha'}] &= -\frac{i}{2\hbar}[Q_{\mathbf{k},\alpha}, P_{\mathbf{k}',\alpha'}] + \frac{i}{2\hbar}[P_{\mathbf{k},\alpha}, Q_{\mathbf{k}',\alpha'}]\\ &= \delta_{\mathbf{k}\mathbf{k}'}\delta_{\alpha\alpha'},\end{aligned} \tag{2.26a}$$

$$[a_{\mathbf{k},\alpha}, a_{\mathbf{k}',\alpha'}] = [a^\dagger_{\mathbf{k},\alpha}, a^\dagger_{\mathbf{k}',\alpha'}] = 0. \tag{2.26b}$$

These communication relations are to be evaluated for operators taken at equal times; for example, $[a_{\mathbf{k},\alpha}, a^\dagger_{\mathbf{k}',\alpha'}]$ actually stands for $[a_{\mathbf{k},\alpha}(t), a^\dagger_{\mathbf{k}',\alpha'}(t)]$.

Before we discuss the physical interpretations of $a_{\mathbf{k},\alpha}$ and $a^\dagger_{\mathbf{k},\alpha}$, it is instructive to study the properties of the operator defined by

$$N_{\mathbf{k},\alpha} = a^\dagger_{\mathbf{k},\alpha}a_{\mathbf{k},\alpha}. \tag{2.27}$$

We have

$$\begin{aligned}[a_{\mathbf{k},\alpha}, N_{\mathbf{k}',\alpha'}] &= a_{\mathbf{k},\alpha}a^\dagger_{\mathbf{k}',\alpha'}a_{\mathbf{k}',\alpha'} - a^\dagger_{\mathbf{k}',\alpha'}a_{\mathbf{k}',\alpha'}a_{\mathbf{k},\alpha}\\ &= [a_{\mathbf{k},\alpha}, a^\dagger_{\mathbf{k}',\alpha'}]a_{\mathbf{k}',\alpha'} - a^\dagger_{\mathbf{k}',\alpha'}[a_{\mathbf{k}',\alpha'}, a_{\mathbf{k},\alpha}]\\ &= \delta_{\mathbf{k}\mathbf{k}'}\delta_{\alpha\alpha'}a_{\mathbf{k},\alpha}.\end{aligned} \tag{2.28}$$

Similarly,

$$[a^\dagger_{\mathbf{k},\alpha}, N_{\mathbf{k}',\alpha}] = -\delta_{\mathbf{k}\mathbf{k}'}\delta_{\alpha\alpha'}a^\dagger_{\mathbf{k},\alpha}. \tag{2.29}$$

Unlike a and $a^\dagger$, the operator N is Hermitian. (In this paragraph and the next we shall suppress the indices $\mathbf{k}$, α. All the relations are valid for given $\mathbf{k}$, α.) The

Hermiticity of N encourages us to consider a normalized eigenvector of the operator N denoted by $|n\rangle$ such that

$$N|n\rangle = n|n\rangle, \tag{2.30}$$

where n is the eigenvalue of N. Because N is Hermitian, n must be real. Now,

$$\begin{aligned} Na^\dagger|n\rangle &= (a^\dagger N + a^\dagger)|n\rangle \\ &= (n+1)a^\dagger|n\rangle, \end{aligned} \tag{2.31}$$

where we have used (2.29). This can be viewed as a new eigenvalue equation in which the eigenvector $a^\dagger|n\rangle$ is shown to have eigenvalue $n+1$. Similarly,

$$Na|n\rangle = (n-1)a|n\rangle. \tag{2.32}$$

The roles of $a^\dagger$ and a are now clear; $a^\dagger(a)$ acting on $|n\rangle$ gives a new eigenvector with eigenvalue increased (decreased) by one. So

$$\begin{aligned} a^\dagger|n\rangle &= c_+|n+1\rangle, \\ a|n\rangle &= c_-|n-1\rangle, \end{aligned} \tag{2.33}$$

where c_+ and c_- are constants. To determine $c_\pm$ we evaluate

$$\begin{aligned} |c_+|^2 &= |c_+|^2\langle n+1|n+1\rangle = \langle a^\dagger n|a^\dagger n\rangle \\ &= \langle n|aa^\dagger|n\rangle = \langle n|N+[a,a^\dagger]|n\rangle \\ &= n+1; \end{aligned} \tag{2.34}$$

$$|c_-|^2 = \langle an|an\rangle = \langle n|a^\dagger a|n\rangle = n. \tag{2.35}$$

The phases of $c_\pm$ are indeterminate; they may be chosen to be zero at $t=0$ by convention. We have, at $t=0$,

$$\begin{aligned} a^\dagger|n\rangle &= \sqrt{n+1}\,|n+1\rangle, \\ a|n\rangle &= \sqrt{n}\,|n-1\rangle. \end{aligned} \tag{2.36}$$

Meanwhile

$$n = \langle n|N|n\rangle = \langle n|a^\dagger a|n\rangle \geq 0, \tag{2.37}$$

since the norm of $a|n\rangle$ must be positive definite. This immediately tells us that n cannot be a noninteger; otherwise, the eigenvalue of $|n\rangle$ could be made to decrease indefinitely as we successively apply a, and n would eventually take negative values, in contradiction to (2.37). On the other hand, if n is a positive integer, successive applications of the operator a proceed as

$$\begin{aligned} a|n\rangle &= \sqrt{n}\,|n-1\rangle\cdots, \qquad a|2\rangle = \sqrt{2}\,|1\rangle, \\ a|1\rangle &= |0\rangle, \qquad a|0\rangle = 0. \end{aligned} \tag{2.38}$$

Note that we obtain a "null vector" (to be distinguished from $|0\rangle$) when a is applied to $|0\rangle$. By applying a to such a null vector, we again obtain a null vector. Hence $n=0$ is the lowest possible eigenvalue of the operator N.

Explicit matrix representations of a, $a^\dagger$, and N consistent with the commutation relations (2.26), (2.28), and (2.29) can be written as follows:

$$a = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & 0 & 0 & \cdots \\ 0 & 0 & 0 & \sqrt{3} & 0 & \cdots \\ 0 & 0 & 0 & 0 & \sqrt{4} & \cdots \\ \vdots & & & & & \end{pmatrix}, \qquad a^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & 0 & \cdots \\ \vdots & & & & & \end{pmatrix},$$

$$N = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 2 & 0 & \cdots \\ 0 & 0 & 0 & 3 & \cdots \\ \vdots & & & & \end{pmatrix}. \tag{2.39}$$

They are assumed to act on a column vector represented by

$$|n\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, \tag{2.40}$$

where only the $(n+1)$-entry is different from zero.

Photon states. The algebra developed above can be applied to a physical situation in which the number of photons with given momentum and polarization is increased or decreased. The wave vector $\mathbf{k}$ will later be identified with the photon momentum divided by $\hbar$, and α will be shown to represent the polarization state of the photon. We interpret an eigenvector of $N_{\mathbf{k},\alpha}$ as the state vector for a state with a definite number of photons in state $(\mathbf{k}, \alpha)$. To represent a situation in which there are many types of photons with different sets of $(\mathbf{k}, \alpha)$, we consider the direct product of eigenvectors as follows:

$$|n_{\mathbf{k}_1,\alpha_1}, n_{\mathbf{k}_2,\alpha_2}, \ldots, n_{\mathbf{k}_l,\alpha_l}, \ldots\rangle = |n_{\mathbf{k}_1,\alpha_1}\rangle |n_{\mathbf{k}_2,\alpha_2}\rangle \cdots |n_{\mathbf{k}_l,\alpha_l}\rangle \cdots \tag{2.41}$$

This state vector corresponds to the physical situation in which there are $n_{\mathbf{k}_1,\alpha_1}$ photons present in state $(\mathbf{k}_1, \alpha_1)$, $n_{\mathbf{k}_2,\alpha_2}$ photons in state $(\mathbf{k}_2, \alpha_2)$, etc. The number $n_{\mathbf{k},\alpha}$ is called the *occupation number* for state $(\mathbf{k}, \alpha)$.

As an example, the state represented by

$$|0\rangle = |0_{\mathbf{k}_1,\alpha_1}\rangle |0_{\mathbf{k}_2,\alpha_2}\rangle \cdots |0_{\mathbf{k}_i,\alpha_i}\rangle \cdots \tag{2.42}$$

has the property that if $a_{\mathbf{k},\alpha}$ is applied to it, then we obtain a null vector for any $(\mathbf{k}, \alpha)$. Hence the eigenvalue of $N_{\mathbf{k},\alpha} = a^\dagger_{\mathbf{k},\alpha} a_{\mathbf{k},\alpha}$ is zero for any $(\mathbf{k}, \alpha)$. Therefore the state corresponding to (2.42) is called the *vacuum state*. A single-photon state with definite $(\mathbf{k}, \alpha)$ is represented by

$$a^\dagger_{\mathbf{k},\alpha} |0\rangle, \tag{2.43}$$

since the eigenvalue of $N_{\mathbf{k},\alpha}$ is one (cf. Eq. 2.31). A two-photon state is represented by the normalized eigenvector

$$(1/\sqrt{2}) a^\dagger_{\mathbf{k},\alpha} a^\dagger_{\mathbf{k},\alpha} |0\rangle \tag{2.44}$$

when the two photons are in the *same* state, and

$$a^\dagger_{\mathbf{k},\alpha} a^\dagger_{\mathbf{k}',\alpha'} |0\rangle. \tag{2.45}$$

when the two photons are in *different* states. More generally,

$$|n_{\mathbf{k}_1,\alpha_1}, n_{\mathbf{k}_2,\alpha_2}, \ldots\rangle = \prod_{\mathbf{k}_i,\alpha_i} \frac{(a^\dagger_{\mathbf{k}_i,\alpha_i})^{n_{\mathbf{k}_i,\alpha_i}}}{\sqrt{n_{\mathbf{k}_i,\alpha_i}!}} |0\rangle. \tag{2.46}$$

Note that when $a^\dagger_{\mathbf{k}_i,\alpha_i}$ is applied to the most general state vector, we obtain

$$a^\dagger_{\mathbf{k}_i,\alpha_i} |n_{\mathbf{k}_1,\alpha_1}, n_{\mathbf{k}_2,\alpha_2}, \ldots, n_{\mathbf{k}_i,\alpha_i}, \ldots\rangle$$
$$= \sqrt{n_{\mathbf{k}_i,\alpha_i} + 1}\, |n_{\mathbf{k}_1,\alpha_1}, n_{\mathbf{k}_2,\alpha_2}, \ldots, n_{\mathbf{k}_i,\alpha_i} + 1, \ldots\rangle. \tag{2.47}$$

Thus $a^\dagger_{\mathbf{k}_i,\alpha_i}$ has the property of creating an additional photon in state $(\mathbf{k}_i, \alpha_i)$, leaving the occupuation numbers of states other than $(\mathbf{k}_i, \alpha_i)$ unchanged. For this reason $a^\dagger_{\mathbf{k},\alpha}$ is called the creation operator for a photon in state $\mathbf{k}, \alpha$. Similarly, $a_{\mathbf{k},\alpha}$ can be interpreted as the annihilation or destruction operator for a photon in state $\mathbf{k}, \alpha$. In contrast, $N_{\mathbf{k},\alpha}$, being diagonal, does not change the occupation number of photons; it simply gives as its eigenvalue the number of photons in state $\mathbf{k}, \alpha$. We might say that the three operators $a^\dagger_{\mathbf{k},\alpha}$, $a_{\mathbf{k},\alpha}$, and $N_{\mathbf{k},\alpha}$ correspond respectively to the Creator (Brahma), the Destroyer (Siva), and the Preserver (Vishnu) in Hindu mythology.

Our formalism is capable of describing a physical situation in which the number of photons in a given state is unrestricted. Moreover, any many-photon system we can construct is necessarily symmetric under interchange of any pair of labels. For instance, the two-photon state (2.45) is evidently symmetric under interchange $(\mathbf{k}, \alpha) \longleftrightarrow (\mathbf{k}', \alpha')$ since

$$a^\dagger_{\mathbf{k},\alpha} a^\dagger_{\mathbf{k}',\alpha'} |0\rangle = a^\dagger_{\mathbf{k}',\alpha'} a^\dagger_{\mathbf{k},\alpha} |0\rangle \tag{2.48}$$

because of the commutation rule. Thus the state vectors we obtain by applying creation operators to $|0\rangle$ are automatically consistent with *Bose-Einstein statistics*. With essentially no modifications our formalism can be applied to physical states made up of indetical particles other than photons as long as they obey Bose-Einstein statistics.

Fermion operators. There exist, in nature, particles that do not obey Bose-Einstein statistics but rather obey *Fermi-Dirac statistics*—electron, muons, protons, etc. For such particles the formalism we have developed is obviously inadequate. We must somehow incorporate the Pauli exclusion principle. This can be done. In 1928 P. Jordan and E. P. Wigner proposed a formalism in which we again consider the operators $b_r^\dagger$ and b_r, but they now satisfy the the "anticommutation relations"

$$\{b_r, b_{r'}^\dagger\} = \delta_{rr'}, \qquad \{b_r, b_{r'}\} = \{b_r^\dagger, b_{r'}^\dagger\} = 0, \tag{2.49}$$

where the term in braces is defined by

$$\{A, B\} = AB + BA. \tag{2.50}$$

The operators $b_r^\dagger$ and b_r are again interpreted as the creation and annihilation operators, and the index r provides a collective description of the momentum state, the spin state, and according to the Dirac hole theory (to be discussed in Sections 3–9 and 3–10), the sign of the energy as well. A single-particle state can be constructed just as before:

$$|1_r\rangle = b_r^\dagger |0\rangle, \tag{2.51}$$

but, since

$$b_r^\dagger b_r^\dagger |0\rangle = \tfrac{1}{2}\{b_r^\dagger, b_r^\dagger\}|0\rangle = 0 \tag{2.52}$$

according to (2.49), we cannot put two particles in the same state. This is just what is needed if the electrons are to satisfy the Pauli exclusion principle. However, if $r \neq r'$, then we can construct a two-particle state

$$b_r^\dagger b_{r'}^\dagger |0\rangle = -b_{r'}^\dagger b_r^\dagger |0\rangle. \tag{2.53}$$

Note that it is necessarily antisymmetric under interchange $r \longleftrightarrow r'$, in conformity with Fermi-Dirac statistics. We define a Hermitian operator N_r by

$$N_r = b_r^\dagger b_r \tag{2.54}$$

just as before. When $b_r^\dagger$ and b_r are interpreted as the creation and annihilation operators, it is natural to regard N_r as the occupation number operator:

$$\begin{aligned} N_r |0\rangle &= b_r^\dagger b_r |0\rangle = 0; \\ N_r b_r^\dagger |0\rangle &= b_r^\dagger (1 - b_r^\dagger b_r)|0\rangle = b_r^\dagger |0\rangle. \end{aligned} \tag{2.55}$$

In general, N_r has the property

$$N_r^2 = b_r^\dagger b_r b_r^\dagger b_r = b_r^\dagger(-b_r^\dagger b_r + 1)b_r = N_r. \tag{2.56}$$

Hence

$$N_r(N_r - 1) = 0, \tag{2.57}$$

which means that the eigenvalue of N_r is either zero or one. Physically speaking, state r is either unoccupied or occupied by just one electron. Explicit matrix representations of b_r, $b_r^\dagger$, and N_r consistent with the anticommunication relations are not difficult to find:

$$b_r = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad b_r^\dagger = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad N_r = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.58}$$

They act on the following column vectors:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.59}$$

Although the algebra of b and $b^\dagger$ is similar to that of a and $a^\dagger$, it is important to note that b and $b^\dagger$ cannot be written as linear combinations of P and Q, satisfying the commutation relations (2.24). This point can be shown to be related to the fact that there is no classically measurable field that corresponds to the quantized fermion field. In this chapter we shall not say anything more about the fermion operators, but we shall come back to this subject in Section 3–10.

2–3. QUANTIZED RADIATION FIELD

Photons as quantum-mechanical excitations of the radiation field. The Fourier coefficients in the expansion of the classical radiation field must be replaced by the corresponding annihilation and creation operators if the canonical variables of the radiation oscillator are to be interpreted as noncommutative quantum-mechanical operators. With the substitutions

$$c_{\mathbf{k},\alpha}(t) \to c\sqrt{\hbar/2\omega}\, a_{\mathbf{k},\alpha}(t) \quad \text{and} \quad c^*_{\mathbf{k},\alpha}(t) \to c\sqrt{\hbar/2\omega}\, a^\dagger_{\mathbf{k},\alpha}(t),$$

we have‡

$$\mathbf{A}(\mathbf{x},t) = (1/\sqrt{V}) \sum_{\mathbf{k}} \sum_{\alpha} c\sqrt{\hbar/2\omega}\,[a_{\mathbf{k},\alpha}(t)\boldsymbol{\epsilon}^{(\alpha)} e^{i\mathbf{k}\cdot\mathbf{x}} + a^\dagger_{\mathbf{k},\alpha}(t)\boldsymbol{\epsilon}^{(\alpha)} e^{-i\mathbf{k}\cdot\mathbf{x}}]. \tag{2.60}$$

Although this expansion is similar in appearance to (2.14), the meaning of $\mathbf{A}$ is very different. The $\mathbf{A}$ of (2.14) is a classical function with three components defined at each space-time point. In contrast, the $\mathbf{A}$ of (2.60) is an operator that acts on state vectors in occupation number space along the lines discussed in the previous section. Note, however, that it is parametrized by $\mathbf{x}$ and t just as a classical field. Such an operator is called a *field operator* or a *quantized field.*

The *Hamiltonian operator* of the quantized radiation field can be taken as

$$H = \tfrac{1}{2} \int (\mathbf{B}\cdot\mathbf{B} + \mathbf{E}\cdot\mathbf{E}) d^3x. \tag{2.61}$$

It can be evaluated using equations such as (2.16) and (2.17) just as before. This time, however, we must be careful with the order of $a_{\mathbf{k},\alpha}$ and $a^\dagger_{\mathbf{k},\alpha}$ since they are no longer mere numbers. We obtain

$$\begin{aligned} H &= \tfrac{1}{2} \sum_{\mathbf{k}} \sum_{\alpha} \hbar\omega (a^\dagger_{\mathbf{k},\alpha} a_{\mathbf{k},\alpha} + a_{\mathbf{k},\alpha} a^\dagger_{\mathbf{k},\alpha}) \\ &= \sum_{\mathbf{k}} \sum_{\alpha} (N_{\mathbf{k},\alpha} + \tfrac{1}{2})\hbar\omega, \end{aligned} \tag{2.62}$$

‡In Gaussian unrationalized units, $c\sqrt{\hbar/2\omega}$ should be replaced by $c\sqrt{2\pi\hbar/\omega}$ because the expression for the total energy has $1/8\pi$ in place of $1/2$ (cf. Eq. 2.16).

with $\omega = |\mathbf{k}|c$. Since the absolute energy scale is arbitrary, we may use the energy scale in which the energy of the vacuum state is zero,

$$H|0\rangle = 0. \tag{2.63}$$

This amounts to subtracting $\sum \sum \hbar\omega/2$ from (2.62), so that (2.62) is replaced by

$$H = \sum_{\mathbf{k}} \sum_{\alpha} \hbar\omega N_{\mathbf{k},\alpha}. \tag{2.64}$$

The Hamiltonian operator H acting on a many-photon state gives

$$H|n_{\mathbf{k}_1,\alpha_1}, n_{\mathbf{k}_2,\alpha_2}, \ldots\rangle = \sum_i n_{\mathbf{k}_i,\alpha_i}\hbar\omega_i |n_{\mathbf{k}_1,\alpha_1}, n_{\mathbf{k}_2,\alpha_2}, \ldots\rangle. \tag{2.65}$$

The total momentum of the radiation field is given in classical electrodynamics as the space integral of the Poynting vector $(\mathbf{E}\times\mathbf{B})/c$.‡ Using an operator expression identical in appearance to the total momentum of the field in classical electrodynamics, we obtain the momentum operator as follows:

$$\begin{aligned}\mathbf{P} = (1/c)\int (\mathbf{E}\times\mathbf{B})d^3x &= \sum_{\mathbf{k}}\sum_{\alpha} \tfrac{1}{2}\hbar\mathbf{k}(a^\dagger_{\mathbf{k},\alpha}a_{\mathbf{k},\alpha} + a_{\mathbf{k},\alpha}a^\dagger_{\mathbf{k},\alpha}) \\ &= \sum_{\mathbf{k}}\sum_{\alpha} \hbar\mathbf{k}(N_{\mathbf{k},\alpha} + \tfrac{1}{2}),\end{aligned} \tag{2.66}$$

where we have used relations such as

$$\frac{c^2\hbar}{2\sqrt{\omega\omega'}}\left(\frac{1}{c}\right)a_{\mathbf{k},\alpha}\left(\frac{\omega'}{c}\right)a^\dagger_{\mathbf{k}',\alpha'}\boldsymbol{\epsilon}^{(\alpha)}\times(\mathbf{k}'\times\boldsymbol{\epsilon}^{(\alpha')})\frac{1}{V}\int e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}}d^3x$$

$$= \frac{\hbar}{2}a_{\mathbf{k},\alpha}a^\dagger_{\mathbf{k},\alpha}\mathbf{k}\delta_{\mathbf{k}\mathbf{k}'}\delta_{\alpha\alpha'}. \tag{2.67}$$

(*Note*: $\boldsymbol{\epsilon}^{(1)}\times(\mathbf{k}\times\boldsymbol{\epsilon}^{(2)}) = 0$, $\boldsymbol{\epsilon}^{(1)}\times(\mathbf{k}\times\boldsymbol{\epsilon}^{(1)}) = \mathbf{k}$, etc.) The $\frac{1}{2}$ that appears in (2.66) can be dropped because when we sum over the allowed $\mathbf{k}$, $\hbar\mathbf{k}$ cancels with $-\hbar\mathbf{k}$. Hence

$$\mathbf{P} = \sum_{\mathbf{k}}\sum_{\alpha} \hbar\mathbf{k}N_{\mathbf{k},\alpha}. \tag{2.68}$$

Let us consider the effect of H and $\mathbf{P}$ on a single-photon state:

$$Ha^\dagger_{\mathbf{k},\alpha}|0\rangle = \hbar\omega a^\dagger_{\mathbf{k},\alpha}|0\rangle, \qquad \mathbf{P}a^\dagger_{\mathbf{k},\alpha}|0\rangle = \hbar\mathbf{k}a^\dagger_{\mathbf{k},\alpha}|0\rangle. \tag{2.69}$$

We see that $\hbar\omega = \hbar|\mathbf{k}|c$ and $\hbar\mathbf{k}$ are respectively the energy and the momentum of the photon. The mass of the photon is given by

$$\begin{aligned}(\text{mass})^2 &= (1/c^4)(E^2 - |\mathbf{p}|^2c^2) \\ &= \frac{1}{c^4}[(\hbar\omega)^2 - (\hbar|\mathbf{k}|c)^2] \\ &= 0.\end{aligned} \tag{2.70}$$

The photon state is characterized not just by its momentum but also by the polarization vector $\boldsymbol{\epsilon}^{(\alpha)}$. Since $\boldsymbol{\epsilon}^{(\alpha)}$ transforms like a vector, the general theory of angular momentum encourages us to associate with it one unit of angular momentum. This is what is meant by the statement that the photon has one unit of *spin*

‡This expression can also be obtained by evaluating the 4-k component of the energy momentum tensor of the transverse electromagnetic field (cf. Problem 1–1).

angular momentum.‡ To find spin components we first consider

$$\boldsymbol{\epsilon}^{(\pm)} = \mp(1/\sqrt{2})(\boldsymbol{\epsilon}^{(1)} \pm i\boldsymbol{\epsilon}^{(2)}), \tag{2.71}$$

which are called *circular polarization vectors.* Under an infinitesimal rotation around the propagation direction $\mathbf{k}$ by an amount $\delta\phi$, the circular polarization vectors are changed by

$$\begin{aligned}\delta\boldsymbol{\epsilon}^{(\pm)} &= \mp(\delta\phi/\sqrt{2})(\boldsymbol{\epsilon}^{(2)} \mp i\boldsymbol{\epsilon}^{(1)}) \\ &= \mp i\delta\phi\boldsymbol{\epsilon}^{(\pm)}.\end{aligned} \tag{2.72}$$

Hence we associate with $\boldsymbol{\epsilon}^{(\pm)}$ the spin component $m = \pm 1$ where the quantization axis has been chosen in the propagation direction. If $\boldsymbol{\epsilon}^{(\alpha)}$ were along $\mathbf{k}$, we would associate $m = 0$ (since $\mathbf{k}$ is unchanged under an infinitesimal rotation about $\mathbf{k}$); however, the $m = 0$ state is missing in the expansion of $\mathbf{A}$ because of the transversality condition $\mathbf{k}\cdot\boldsymbol{\epsilon}^{(\alpha)} = 0$. In other words, the photon spin is either parallel or antiparallel to the propagation direction. We note that the absence of the $m = 0$ state has an invariant meaning only for a particle whose mass is strictly zero. If the photon mass is *different* from zero, we can perform a Lorentz transformation such that in the new frame the photon is at rest; in such a frame the photon spin is "parallel to nothing."

The description of the polarization state with $\boldsymbol{\epsilon}^{\pm}$ as the base vectors is called the *circular polarization representation* in contrast to the linear polarization representation based on $\boldsymbol{\epsilon}^{(1)}$ and $\boldsymbol{\epsilon}^{(2)}$. The orthogonality relations for $\boldsymbol{\epsilon}^{(\pm)}$ are

$$\begin{aligned}\boldsymbol{\epsilon}^{(\pm)}\cdot\boldsymbol{\epsilon}^{(\pm)*} &= -\boldsymbol{\epsilon}^{(\pm)}\cdot\boldsymbol{\epsilon}^{(\mp)} = 1, \\ \boldsymbol{\epsilon}^{(\pm)}\cdot\boldsymbol{\epsilon}^{(\mp)*} &= -\boldsymbol{\epsilon}^{(\pm)}\cdot\boldsymbol{\epsilon}^{(\pm)} = 0,\end{aligned} \tag{2.73a}$$

$$\mathbf{k}\cdot\boldsymbol{\epsilon}^{(\pm)} = 0. \tag{2.73b}$$

We could have started with an expansion of $\mathbf{A}$ with $\boldsymbol{\epsilon}^{(\pm)}$ in place of $\boldsymbol{\epsilon}^{(1)}$ and $\boldsymbol{\epsilon}^{(2)}$. A single-photon state with definite circular polarization can be constructed by applying the creation operator

$$a^{\dagger}_{\mathbf{k},\pm} = \mp\frac{1}{\sqrt{2}}(a^{\dagger}_{\mathbf{k},1} \mp i a^{\dagger}_{\mathbf{k},2}) \tag{2.74}$$

to the vacuum state. Conversely, $a^{\dagger}_{\mathbf{k},\alpha}|0\rangle$ with $\alpha = 1, 2$ and $\mathbf{k}$ in the z-direction can be regarded as a 50/50 mixture of the $m = \mp 1$ and $m = -1$ state.

To sum up, the quantum postulate applied to the canonical variables of the radiation oscillator naturally leads to the idea that the quantum-mechanical excitations of the radiation field can be regarded as particles of *mass* zero and *spin* one. It is a general feature of the quantum theory of fields that *with every field we associate a particle of definite mass and spin.* The arguments we have presented can be repeated for the quantization of other fields with essentially no modifications.

---

‡For a further discussion of this point, see Problem 2–2. The spin angular momentum of the radiation field can also be discussed from the point of view of the total angular-momentum operator $(1/c)\int[\mathbf{x} \times (\mathbf{E} \times \mathbf{B})]d^3x$ which can be decomposed into the orbital and the spin part. See, for example, Wentzel (1949), p. 123 and Messiah (1962), pp. 1022–1024, pp. 1032–1034.

To study the time development of the quantized radiation field, let us first note that $a_{\mathbf{k},\alpha}$ and $a^\dagger_{\mathbf{k},\alpha}$ are time-*dependent* operators satisfying the Heisenberg equation of motion:

$$\begin{aligned}\dot{a}_{\mathbf{k},\alpha} &= (i/\hbar)[H, a_{\mathbf{k},\alpha}] \\ &= (i/\hbar)\sum_{\mathbf{k}',\alpha'}[\hbar\omega' N_{\mathbf{k}',\alpha'}, a_{\mathbf{k},\alpha}] \\ &= -i\omega a_{\mathbf{k},\alpha}; \end{aligned} \tag{2.75}$$

$$\begin{aligned}\ddot{a}_{\mathbf{k},\alpha} &= (i/\hbar)[H, \dot{a}_{\mathbf{k},\alpha}] = (i/\hbar)[H, -i\omega a_{\mathbf{k},\alpha}] \\ &= -\omega^2 a_{\mathbf{k},\alpha}. \end{aligned} \tag{2.76}$$

These are identical in appearance with the differential equations satisfied by the Fourier coefficients $c_{\mathbf{k},\alpha}(t)$ of the classical radiation field. Similarly,

$$\dot{a}^\dagger_{\mathbf{k},\alpha} = i\omega a^\dagger_{\mathbf{k},\alpha}, \qquad \ddot{a}^\dagger_{\mathbf{k},\alpha} = -\omega^2 a^\dagger_{\mathbf{k},\alpha}. \tag{2.77}$$

From (2.76) and (2.77) it is evident that the quantized field **A** satisfies the same wave equation (2.7) as does the classical field. Integrating (2.75) we obtain the explicit time dependence for the annihilation and the creation operator:

$$a_{\mathbf{k},\alpha} = a_{\mathbf{k},\alpha}(0)e^{-i\omega t}, \qquad a^\dagger_{\mathbf{k},\alpha} = a^\dagger_{\mathbf{k},\alpha}(0)e^{i\omega t}. \tag{2.78}$$

Finally, we have

$$\mathbf{A}(\mathbf{x}, t) = \frac{1}{\sqrt{V}}\sum_{\mathbf{k}}\sum_{\alpha} c\sqrt{\frac{\hbar}{2\omega}}[a_{\mathbf{k},\alpha}(0)\boldsymbol{\epsilon}^{(\alpha)}e^{i\mathbf{k}\cdot\mathbf{x}-i\omega t} + a^\dagger_{\mathbf{k},\alpha}(0)\boldsymbol{\epsilon}^{(\alpha)}e^{-i\mathbf{k}\cdot\mathbf{x}+i\omega t}]. \tag{2.79}$$

Note that the quantized field operator **A** is Hermitian while the classical field **A** is real. It is important to note that the **x** and t that appear in the quantized field $\mathbf{A}(\mathbf{x}, t)$ are not quantum-mechanical variables but just parameters on which the field operator depends. In particular, **x** and t should not be regarded as the space-time coordinates of the photon.

Fluctuations and the uncertainty relations. We are now in a position to discuss some of the peculiarities arising from the quantum nature of the radiation field. First, let us note that neither the individual occupation number operator $N_{\mathbf{k},\alpha}$ nor the total number operator defined by

$$N = \sum_{\mathbf{k}}\sum_{\alpha} N_{\mathbf{k},\alpha} = \sum_{\mathbf{k}}\sum_{\alpha} a^\dagger_{\mathbf{k},\alpha}a_{\mathbf{k},\alpha} \tag{2.80}$$

commutes with **A**, **E**, and **B** because of (2.28) and (2.29). As is well known, in quantum mechanics the observables corresponding to noncommutable operators cannot be simultaneously determined to arbitrary degrees of accuracy. In the present case, if the number of photons is approximately fixed, there must necessarily be uncertainties in the field strengths.‡ Such a fluctuation behavior is expected even for the vacuum state. To check this explicitly let us take the

‡In general, by "uncertainty" we mean the square root of the deviation of the expectation value of the square of the operator in question from the square of the expectation value of the operator, that is, $\Delta q = \sqrt{\langle q^2\rangle - \langle q\rangle^2}$.

electric field operator $\mathbf{E} = -(1/c)\partial\mathbf{A}/\partial t$. Although the vacuum expectation value of the electric field $\langle 0|\mathbf{E}|0\rangle$ vanishes (because $a_{\mathbf{k},\alpha}|0\rangle = 0$) as expected from symmetry considerations, the mean square fluctuation of the electric field can be readily shown to be infinite:

$$\langle 0|\mathbf{E}\cdot\mathbf{E}|0\rangle - |\langle 0|\mathbf{E}|0\rangle|^2 = \langle 0|\mathbf{E}\cdot\mathbf{E}|0\rangle = \infty. \tag{2.81}$$

This illustrates our assertion that if the occupation number is fixed (zero in this case), then the field strength is completely uncertain. On the other hand, since what we measure by a test body is the field strength averaged over some region in space, it may be more realistic to consider the average field operator about some point, for example the origin, defined by

$$\bar{\mathbf{E}} = (1/\Delta V)\int_{\Delta V} \mathbf{E}\, d^3x, \tag{2.82}$$

where ΔV is a small volume containing the point in question. We can then prove

$$\langle 0|\bar{\mathbf{E}}\cdot\bar{\mathbf{E}}|0\rangle \sim \hbar c/(\Delta l)^4, \tag{2.83}$$

where Δl is the linear dimension of the volume ΔV.‡ Equation (2.83) characterizes the fluctuation in the electric field when no photons are present. In general, if the set of occupation numbers is approximately fixed, then the electric or magnetic field has no precise value but exhibits fluctuations about a certain average value.

Another peculiar feature of the quantized radiation field is that the components of $\mathbf{E}$ and $\mathbf{B}$ do not necessarily commute. For instance, we have§

$$[E_x(\mathbf{x}, t), B_y(\mathbf{x}', t)] = ic\hbar\frac{\partial}{\partial z}\delta^3(\mathbf{x} - \mathbf{x}'). \tag{2.84}$$

It follows that we cannot simultaneously determine $\mathbf{E}$ and $\mathbf{B}$ at the same point in space to arbitrary degrees of accuracy. However, it is possible to show that

$$[E_x(\mathbf{x}, t), B_y(\mathbf{x}', t')] = 0 \qquad \text{for} \qquad (\mathbf{x} - \mathbf{x}')^2 - (1/c^2)(t - t')^2 \neq 0. \tag{2.85}$$

Thus there is no interference between measurements of field strengths performed at two space-time points separated by a distance that cannot be connected by a light signal. This is in comformity with the principle of causality in the special theory of relativity. The commutation relations (2.84), (2.85), and other similar relations were first obtained by P. Jordan, W. Heisenberg, and W. Pauli.

Using the commutation relation (2.84) we can write an uncertainty relation that involves $\Delta\mathbf{E}$ and $\Delta\mathbf{B}$. However, before we interpret such a relation physically we must first define what is meant in the quantum theory by the measurement of field strength. This would require a rather careful discussion based on a 1933 paper by N. Bohr and L. Rosenfeld. Consequently we shall not discuss this subject any further. An excellent treatment of it can be found in Heitler's book.‖

---

‡We shall not prove (2.83) since in Problem 2–3 the reader is asked to prove a similar relation for the scalar field.

§Again we shall not prove this relation because Problem 2–3 is concerned with a similar relation for the scalar field.

‖Heitler (1954), pp. 76–86.

We know from classical optics that a localized or converging light beam can be constructed by superposing various plane waves with definite phase relations. For this reason we wish to consider the phase operator $\phi_{\mathbf{k},\alpha}$ associated with given $\mathbf{k}$, α such that the observable corresponding to it is the phase of the plane wave in question in the sense of classical optics. This we can do by setting

$$\begin{aligned} a_{\mathbf{k},\alpha} &= e^{i(\phi_{\mathbf{k},\alpha}-\omega t)}\sqrt{N_{\mathbf{k},\alpha}}, \\ a^{\dagger}_{\mathbf{k},\alpha} &= \sqrt{N_{\mathbf{k},\alpha}}\,e^{-i(\phi_{\mathbf{k},\alpha}-\omega t)}, \end{aligned} \tag{2.86}$$

where the operator $\sqrt{N_{\mathbf{k},\alpha}}$ satisfies the property $(\sqrt{N_{\mathbf{k},\alpha}})^2 = N_{\mathbf{k},\alpha}$. The quantized radiation field can be written in such a way that the only operators appearing are $\sqrt{N_{\mathbf{k},\alpha}}$ and $\phi_{\mathbf{k},\alpha}$:

$$\begin{aligned} \mathbf{A}(\mathbf{x},t) = \frac{1}{\sqrt{V}}\sum_{\mathbf{k}}\sum_{\alpha} c\sqrt{\frac{\hbar}{2\omega}}\Big(&\boldsymbol{\epsilon}^{(\alpha)}\exp\left[i(\mathbf{k}\cdot\mathbf{x}-\omega t+\phi_{\mathbf{k},\alpha})\right]\sqrt{N_{\mathbf{k},\alpha}} \\ &+\sqrt{N_{\mathbf{k},\alpha}}\boldsymbol{\epsilon}^{(\alpha)}\exp\left[-i(\mathbf{k}\cdot\mathbf{x}-\omega t+\phi_{\mathbf{k},\alpha})\right]\Big), \end{aligned} \tag{2.87}$$

which is of a form convenient for use in discussing the connection with the classical description. The commutation relation for a and $a^{\dagger}$ (for given $\mathbf{k}$, α) can now be written as

$$\begin{aligned} e^{i\phi}Ne^{-i\phi} - N &= 1, \\ e^{i\phi}N - Ne^{i\phi} &= e^{i\phi}, \end{aligned} \tag{2.88}$$

where we have suppressed the subscripts $\mathbf{k}$, α. It is easy to prove by expanding each exponential that the commutation relation (2.88) is satisfied whenever $N\phi - \phi N = i$. This leads to a new uncertainty relation for the corresponding observables:

$$\Delta N\Delta\phi \gtrsim 1. \tag{2.89}$$

For example, if the phase difference of two plane-wave components is given precisely, then we cannot tell the occupation number associated with each of the wave components.

The above uncertainty relation can be used to derive a more familiar-looking uncertainty relation for the light beam. Since the momentum operator is written as the sum of the number operators weighted by $\hbar\mathbf{k}$ (cf. Eq. 2.68) the momentum of the light beam is uncertain whenever there are definite phase relations among individual plane-wave components. Meanwhile a localized wave train can be constructed by superposing plane-wave components with definite phase relations. It then follows that we cannot associate a definite momentum to a well localized beam. For simplicity let us consider a one-dimensional wave train confined to a region whose characteristic dimension is $\sim \Delta x$. Although the actual mathematical description of such a localized wave train is somewhat involved, we expect that the phase of a typical plane wave component must be known qualitatively to an accuracy of $k_x\Delta x$ if the localization is to be possible. The uncertainty relation (2.89) then implies

$$\Delta p_x\,\Delta x \gtrsim \hbar, \tag{2.90}$$

since the momentum uncertainty is of the order of $\Delta N \hbar k_x$. This relation is formally identical to Heisenberg's uncertainty relation for a "particle." However, because of (2.89), a single photon cannot be localized in the same sense that a particle in nonrelativistic quantum mechanics can be localized. For this reason, it is better to regard (2.90) as a relation imposed on the light *beam*, as we have done.

It is amusing to note that Heisenberg's uncertainty principle for a material particle, for example, an electron, can be formulated only when the light beam also satisfies the uncertainty relation (2.90). If there were no restriction on Δx and Δp_x for the light beam, we would be able to determine the position of the electron to an infinite degree of accuracy by a very well-localized beam without transferring an appreciable momentum uncertainty to the electron in an uncontrollable way. Two successive measurements of this kind would determine accurately both the momentum and the position of the electron, in contradiction with Heisenberg's relation,

Validity of the classical description. All these peculiarities of the quantized radiation field could be rather disturbing, especially if we recall that many of the conveniences of our civilization rely upon the validity of classical electrodynamics which, as we have seen, must be modified. For instance, having learned about the quantum fluctuations of the radiation field, how can we listen in peace to an FM broadcast of symphonic music? In nonrelativistic quantum mechanics the classical description is trustworthy whenever the noncommutativity of dynamical variables is unimportant. Likewise, in the quantum theory of radiation, if we could ignore the right-hand side of $[a, a^\dagger] = 1$, then we would return to the classical description. Since the nonvanishing matrix elements of a and $a^\dagger$ are of the order of $\sqrt{n}$ (cf. Eq. 2.36), the occupation number must be large if the classical description is to be valid.

To be more specific, let us compare the vacuum fluctuation of a squared-field operator and the square of the field strength for a classical electromagnetic wave of wavelength $2\pi\lambda\!\!\!^{-}$. When there are no photons of any kind, the square of the average electric field operator has the following expectation value (cf. Eq. 2.83):

$$\langle 0 | \bar{\mathbf{E}} \cdot \bar{\mathbf{E}} | 0 \rangle \sim \hbar c / \lambda\!\!\!^{-4}, \tag{2.91}$$

where the average is taken over a volume $\lambda\!\!\!^{-3}$. Meanwhile, according to classical electrodynamics, the time average of $\mathbf{E}^2$ can be equated with the energy density of the electromagnetic wave. So we set

$$(\mathbf{E}^2)_{\text{average}} = \bar{n}\hbar(c/\lambda\!\!\!^{-}), \tag{2.92}$$

where $\bar{n}$ stands for the number of photons per unit volume. For the validity of the classical description, purely quantum effects such as (2.91) must be completely negligible in comparison to (2.92). For this we must have

$$\bar{n} \gg 1/\lambda\!\!\!^{-3}. \tag{2.93}$$

In other words, the description of physical phenomena based on classical electrodynamics is reliable when the number of photons per volume $\lambda\!\!\!^{-3}$ is *much greater* than one. As an example, for a Chicago FM station (WFMT), which broadcasts at

98.7 Mc ($\lambda\!\!\!^{-} \approx 48$ cm) with a power of 135,000 watts, the number of photons per volume $\lambda\!\!\!^{-3}$ at a distance five miles from the antenna is about 10^{17}. Thus the classical approximation is an extremely good one.

The historical development of quantum mechanics was guided by an analogy between the electron and the photon both of which were recognized to exhibit the famous wave-particle duality. As Heitler correctly emphasized in his treatise, this similarity can be somewhat misleading. The classical limit of the quantum theory of radiation is achieved when the number of photons becomes so large that the occupation number may as well be regarded as a continuous variable. The space-time development of the classical electromagnetic wave approximates the dynamical behavior of trillions of photons. In contrast, the classical limit of Schrödinger's wave mechanics is the mechanics of a *single* mass point obeying Newton's equation of motion. Thus it was no coincidence that in the very beginning only the *wave nature* of light and the *particle nature* of the electron were apparent.

2–4. EMISSION AND ABSORPTION OF PHOTONS BY ATOMS

Basic matrix elements for emission and absorption. We now have the necessary machinery to deal with the emission and absorption of photons by nonrelativistic atomic electrons. The interaction Hamiltonian between the atomic electrons and the radiation field is assumed to be obtainable from the standard prescription $\mathbf{p} \to \mathbf{p} - e\mathbf{A}/c$, where $\mathbf{A}$ now stands for the quantized radiation field. We have

$$H_{\text{int}} = \sum_i \left[-\frac{e}{2mc}(\mathbf{p}_i \cdot \mathbf{A}(\mathbf{x}_i, t) + \mathbf{A}(\mathbf{x}_i, t) \cdot \mathbf{p}_i) + \frac{e^2}{2mc^2}\mathbf{A}(\mathbf{x}_i, t) \cdot \mathbf{A}(\mathbf{x}_i, t) \right], \tag{2.94}$$

where the summation is over the various atomic electrons that participate in the interaction. The expression $\mathbf{A}(\mathbf{x}_i, t)$ is now a field operator assumed to act on a photon state or a many-photon state at $\mathbf{x}_i$, where $\mathbf{x}_i$ refers to the coordinate of the ith electron. The operator $\mathbf{p}_i$ in $\mathbf{p}_i \cdot \mathbf{A}$ is a differential operator that acts on everything that stands to the right; however, because of the transversality condition $\nabla \cdot \mathbf{A} = 0$, it is legitimate to replace $\mathbf{p}_i \cdot \mathbf{A}$ by $\mathbf{A} \cdot \mathbf{p}_i$. Since the spin magnetic moment is known to interact with the magnetic field, there is an additional interaction of the form

$$H_{\text{int}}^{(\text{spin})} = -\sum_i \frac{e\hbar}{2mc}\boldsymbol{\sigma}_i \cdot [\nabla \times \mathbf{A}(\mathbf{x}, t)]_{\mathbf{x}=\mathbf{x}_i}. \tag{2.95}$$

As we saw in the previous section, $\mathbf{A}(\mathbf{x}, t)$ is a linear combination of the creation and annihilation operators for photons. The question of which particular creation or annihilation operator gives rise to a nonvanishing matrix element depends entirely on the nature of the initial and final states in question.

From the time-dependent perturbation theory (which we will review shortly), it is well known that the transition matrix element for $A \longrightarrow B$ (where A and B may, for example, be atomic states) can be computed to first order in the interaction Hamiltonian H_I by simply taking the matrix element of H_I between A and

B. The H_{int} that appears in (2.94) and (2.95), however, acts not only on atomic states but also on photon states. In a typical process appearing in the quantum theory of radiation the state vector for the initial (or final) state is the *direct product* of the state vector for an atomic state (denoted by A, B, etc.) and the state vector for a single- or a multi-photon state (characterized by $n_{\mathbf{k},\alpha}$). With this point in mind we can still evaluate the transition matrix element to lowest order by taking the matrix element of (2.94) or (2.95) between the initial and final states.

Let us first consider the absorption of a light quantum characterized by $\mathbf{k}, \alpha$. An atom which is initially in state A makes a radiative transition to state B. For simplicity we shall assume that there are only photons of the kind $(\mathbf{k}, \alpha)$ present. If there are $n_{\mathbf{k},\alpha}$ photons in the initial state, then there are $n_{\mathbf{k},\alpha} - 1$ photons in the final state. Although $\mathbf{A}$ contains both $a_{\mathbf{k},\alpha}$ and $a^\dagger_{\mathbf{k},\alpha}$, only $a_{\mathbf{k},\alpha}$ gives rise to a nonvanishing matrix element so long as we are computing the absorption process to lowest order. The quadratic term $\mathbf{A}\cdot\mathbf{A}$ makes no contribution to this process in lowest order since it changes the total number of photons by either 0 or ± 2. So, ignoring the spin magnetic moment interaction, we have

$$\begin{aligned}
&\langle B; n_{\mathbf{k},\alpha} - 1 \,|\, H_{\text{int}} \,|\, A; n_{\mathbf{k},\alpha}\rangle \\
&\quad = -\frac{e}{mc}\langle B; n_{\mathbf{k},\alpha} - 1 \,|\, \sum_i c\sqrt{\frac{\hbar}{2\omega V}}\, a_{\mathbf{k},\alpha}(0) e^{i\mathbf{k}\cdot\mathbf{x}_i - i\omega t}\mathbf{p}_i\cdot\boldsymbol{\epsilon}^{(\alpha)} \,|\, A; n_{\mathbf{k},\alpha}\rangle \\
&\quad = -\frac{e}{m}\sqrt{\frac{n_{\mathbf{k},\alpha}\hbar}{2\omega V}} \sum_i \langle B \,|\, e^{i\mathbf{k}\cdot\mathbf{x}_i}\mathbf{p}_i\cdot\boldsymbol{\epsilon}^{(\alpha)} \,|\, A\rangle e^{-i\omega t},
\end{aligned} \tag{2.96}$$

where we have used (2.36) and (2.79). Note that the annihilation operator for a photon with momentum and polarization different from $\hbar\mathbf{k}$ and $\boldsymbol{\epsilon}^{(\alpha)}$ gives a zero matrix element.

It is instructive to compare this expression with the expression we obtain using the semiclassical theory of radiation in which the vector potential $\mathbf{A}$ is treated classically. Within the framework of the semiclassical theory we may define an equivalent classical vector potential denoted by $\mathbf{A}^{(\text{abs})}$ such that when it is used for an absorption process we obtain a matrix element identical to (2.96). Evidently

$$\mathbf{A}^{(\text{abs})} = \mathbf{A}_0^{(\text{abs})} e^{i\mathbf{k}\cdot\mathbf{x} - i\omega t} \tag{2.97}$$

will do, where

$$\mathbf{A}_0^{(\text{abs})} = c\sqrt{\frac{n_{\mathbf{k},\alpha}\hbar}{2\omega V}}\boldsymbol{\epsilon}^{(\alpha)}. \tag{2.98}$$

According to the semiclassical theory the absorption probability is proportional to the intensity $|\mathbf{A}_0|^2$; in the quantum theory it is proportional to $n_{\mathbf{k},\alpha}$. We expect on general grounds that the semiclassical theory approximates reality for large values of $n_{\mathbf{k},\alpha}$, but since the matrix element (2.96) is linear in $\sqrt{n_{\mathbf{k},\alpha}}$, the semiclassical theory gives the correct answer even for small values of $n_{\mathbf{k},\alpha}$ or for weak radiation.

The above happy accident no longer takes place for emission processes which we will now discuss. This time it is the creation operator $a^\dagger_{\mathbf{k},\alpha}$ that gives rise to a nonvanishing contribution since the photon occupation number is increased by

one. We have (cf. Eqs. 2.36 and 2.79):

$$\langle B; n_{\mathbf{k},\alpha}+1 \,|\, H_{\text{int}} \,|\, A; n_{\mathbf{k},\alpha}\rangle = -\frac{e}{m}\sqrt{\frac{(n_{\mathbf{k},\alpha}+1)\hbar}{2\omega V}} \sum_i \langle B \,|\, e^{-i\mathbf{k}\cdot\mathbf{x}_i}\mathbf{p}_i\cdot\boldsymbol{\epsilon}^{(\alpha)} \,|\, A\rangle e^{i\omega t}. \tag{2.99}$$

If $n_{\mathbf{k},\alpha}$ is very large, the semiclassical treatment based on the complex conjugate of the equivalent potential, (2.97) and (2.98), is adequate since, in practice, there is no difference between $\sqrt{n_{\mathbf{k},\alpha}+1}$ and $\sqrt{n_{\mathbf{k},\alpha}}$.‡ Therefore, for very intense radiation, the problem can again be treated using the semiclassical theory. On the other hand, for small values of $n_{\mathbf{k},\alpha}$ the correspondence with the semiclassical description fails completely. In particular, the transition matrix element for emission need not be zero when there are no photons at all present initially ($n_{\mathbf{k},\alpha}=0$). This accounts for the emission of a photon by an isolated excited atom when there are no electromagnetic waves incident on it, a phenomenon known as "spontaneous emission." In contrast, the emission of a photon in the case where $n_{\mathbf{k},\alpha}\neq 0$ is known as "induced (or stimulated) emission." Note that in the quantum field-theoretic treatment *spontaneous* emission and *induced* emission are discussed on the same footing. The single expression (2.99) suffices for both.

The fact that we can account for spontaneous emission in such a natural way is one of the triumphs of the quantum theory of radiation as opposed to the semiclassical theory of radiation. So long as the electromagnetic field is described by a *classical* potential, the transition matrix element for emission is zero when there is no electromagnetic wave incident on the atom; $\mathbf{A}\cdot\mathbf{p}=0$ when $\mathbf{A}=0$, and that is that. In the *quantum* theory of radiation this is not the case because the matrix element of the field operator $\mathbf{A}(\mathbf{x}, t)$ taken between the vacuum state and the one-photon state does not vanish.

In general, in the classical theory, $\mathbf{A}$ is an externally applied potential that influences the charged particles but is not influenced by them; $\mathbf{A}$ itself does not change as an atom makes a radiative transition. This description is satisfactory, even within the framework of quantum theory, whenever the occupation number is so large that the radiation field can be regarded as an inexhaustible source and sink of photons. It should be clear that, for intense radiation, taking one photon away from the radiation field or adding one photon to it does not make much practical difference. On the other hand, for weak or no incident radiation the classical description runs into difficulty because the change in the radiation field brought about by the emission and absorption of a photon by an atom is quite noticeable.

With the above limitations in mind, we may still define an equivalent classical potential appropriate for emission processes as follows:

$$\mathbf{A}^{(\text{emis})} = \mathbf{A}_0^{(\text{emis})} e^{-i\mathbf{k}\cdot\mathbf{x}+i\omega t}, \tag{2.100}$$

‡We recall that the total time-dependent potential that appears in nonrelativistic quantum mechanics must be Hermitian. In our particular case, the classical vector potential $\mathbf{A}$ must be real anyway.

with

$$\mathbf{A}_0^{(\text{emis})} = c\sqrt{(n_{\mathbf{k},\alpha}+1)\hbar/2\omega V}\,\boldsymbol{\epsilon}^{(\alpha)}. \tag{2.101}$$

By treating (2.100) as the time-dependent vector potential that appears in the Schrödinger equation we can obtain the correct matrix elements (2.99) for emission processes (including spontaneous emission). The fact that $\mathbf{A}^{(\text{emis})}$ is no longer linear in $\sqrt{n_{\mathbf{k},\alpha}}$ and is not quite the complex conjugate of $\mathbf{A}^{(\text{abs})}$ given by (2.97) and (2.98), reflects the failure of the classical concepts.

To sum up, starting with the quantum theory of radiation, we have rigorously derived the following very useful rule.

The emission or absorption of a light quantum by a charged particle is completely equivalent to an interaction of the charged particle with the equivalent unquantized vector potential given below:

$$\begin{aligned} &\text{absorption:} \quad c\sqrt{\frac{n_{\mathbf{k},\alpha}\hbar}{2\omega V}}\boldsymbol{\epsilon}^{(\alpha)}e^{i\mathbf{k}\cdot\mathbf{x}-i\omega t}, \\ &\text{emission:} \quad c\sqrt{\frac{(n_{\mathbf{k},\alpha}+1)\hbar}{2\omega V}}\boldsymbol{\epsilon}^{(\alpha)}e^{-i\mathbf{k}\cdot\mathbf{x}+i\omega t}, \end{aligned} \tag{2.102}$$

where $n_{\mathbf{k},\alpha}$ is the photon occupation number in the initial state.
This simple rule is one of the most important results of this book.‡

Time-dependent perturbation theory. Before we start computing the transition probabilities for various processes let us briefly review the time-dependent perturbation theory developed by Dirac. An atomic wave function ψ can be expanded as

$$\psi = \sum_k c_k(t)u_k(\mathbf{x})e^{-iE_kt/\hbar}, \tag{2.103}$$

where $u_k(\mathbf{x})$ is the energy eigenfunction with energy E_k satisfying

$$H_0u_k(\mathbf{x}) = E_ku_k(\mathbf{x}) \tag{2.104}$$

in the absence of a time-dependent perturbation. The time-dependent Schrödinger equation in the presence of a time-dependent potential $H_I(t)$ is

$$\begin{aligned} (H_0+H_I)\psi &= i\hbar(\partial\psi/\partial t) \\ &= i\hbar\sum_k(\dot{c}_ku_ke^{-iE_kt/\hbar} - i(E_k/\hbar)c_ku_ke^{-iE_kt/\hbar}). \end{aligned} \tag{2.105}$$

In our case, in addition to the kinetic energies of the electrons the unperturbed Hamiltonian H_0 contains the Coulomb interactions between the electrons and the nucleus, whereas $H_I(t)$ accounts for the interaction of the atomic electrons with the equivalent vector potential (2.102). Using (2.104), we have

$$\sum_k H_Ic_ku_ke^{-iE_kt/\hbar} = i\hbar\sum_l \dot{c}_lu_le^{-iE_lt/\hbar}. \tag{2.106}$$

---

‡We again emphasize that in Gaussian unrationalized units $c\sqrt{\hbar/2\omega}$ is to be replaced by $c\sqrt{2\pi\hbar/\omega}$.

Multiplying $u_m^* e^{iE_m t/\hbar}$ and integrating over the space coordinates, we obtain the differential equation

$$\dot{c}_m = \sum_k (1/i\hbar)\langle m|H_I(t)|k\rangle e^{i(E_m - E_k)t/\hbar} c_k(t). \tag{2.107}$$

Suppose only state l is populated when the time-dependent perturbation is turned on at $t = 0$; then

$$c_k(0) = \delta_{kl}. \tag{2.108}$$

We may then approximate c_m by

$$c_m^{(1)}(t) = \frac{1}{i\hbar}\int_0^t dt' \langle m|H_I(t')|l\rangle e^{i(E_m - E_l)t'/\hbar}. \tag{2.109}$$

If $c_m^{(1)}$ vanishes for some reason or if a better approximation is called for, we use

$$c_m \simeq c_m^{(1)} + c_m^{(2)}, \tag{2.110}$$

where

$$c_m^{(2)}(t) = \frac{1}{i\hbar}\sum_n \int_0^t dt'' \langle m|H_I(t'')|n\rangle e^{i(E_m - E_n)t''/\hbar} c_n^{(1)}(t'')$$

$$= \frac{1}{(i\hbar)^2}\sum_n \int_0^t dt'' \int_0^{t''} dt' \langle m|H_I(t'')|n\rangle e^{i(E_m - E_n)t''/\hbar}\langle n|H_I(t')|l\rangle e^{i(E_n - E_l)t'/\hbar}, \tag{2.111}$$

and so on.

When we are dealing with the emission and absorption of a photon by an atom, we may work with just $c_m^{(1)}$. The time-dependent potential H_I can be written as

$$H_I(t) = H_I' e^{\mp i\omega t} \qquad \text{for} \quad \left\{\begin{matrix}\text{absorption}\\ \text{emission}\end{matrix}\right\}, \tag{2.112}$$

where H_I' is a time-independent operator. Hence

$$c_m^{(1)} = \frac{1}{i\hbar}\langle m|H_I'|l\rangle \int_0^t dt' e^{i(E_m - E_l \mp \hbar\omega)t'/\hbar}. \tag{2.113}$$

We can readily perform the time integration and obtain

$$|c_m^{(1)}(t)|^2 = (2\pi/\hbar)|\langle m|H_I'|l\rangle|^2 t\,\delta(E_m - E_l \mp \hbar\omega), \tag{2.114}$$

where we have used

$$\lim_{\alpha\to\infty}\frac{1}{\pi}\frac{\sin^2 \alpha x}{\alpha x^2} = \delta(x). \tag{2.115}$$

Note that the transition probability per unit time is $|c_m^{(1)}(t)|^2/t$, which is independent of t.

Let us now apply our formalism to emission processes. The photon states allowed by the periodic boundary conditions form a continuous energy spectrum as the normalization volume becomes infinite (cf. Eq. 2.11). For a photon emitted into a solid angle element $d\Omega$, the number of allowed states in an energy interval $[\hbar\omega, \hbar(\omega + d\omega)]$ can be written as $\rho_{\hbar\omega, d\Omega}\, d(\hbar\omega)$, where

$$\rho_{\hbar\omega, d\Omega} = \frac{V|\mathbf{k}|^2}{(2\pi)^3}\frac{d|\mathbf{k}|}{d(\hbar\omega)}d\Omega = \frac{V\omega^2}{(2\pi)^3}\frac{d\Omega}{\hbar c^3}. \tag{2.116}$$

So for the transition probability per unit time into a solid angle element $d\Omega$ we obtain the famous *Golden Rule*

$$w_{d\Omega} = \int (|c_m^{(1)}|^2/t)\rho_{\hbar\omega,\,d\Omega}\,d(\hbar\omega)$$
$$= (2\pi/\hbar)\,|\langle m|\,H_I'\,|\,l\rangle|^2\rho_{\hbar\omega,\,d\Omega}, \tag{2.117}$$

where $\hbar\omega$ must satisfy

$$E_m - E_l + \hbar\omega = 0. \tag{2.118}$$

Spontaneous emission in the dipole approximation. In the case of spontaneous emission, an atomic state A makes a radiative transition to a state B in the absence of any incident electromagnetic wave. The matrix element $\langle B\,|H_I'\,|\,A\rangle$ in this case is just (2.99) with $e^{i\omega t}$ omitted and $n_{\mathbf{k},\alpha}$ set equal to zero. Hence for $w_{d\Omega}$ we obtain

$$w_{d\Omega} = \frac{2\pi}{\hbar}\frac{e^2\hbar}{2m^2\omega V}\Big|\sum_i \langle B|\,e^{-i\mathbf{k}\cdot\mathbf{x}_i}\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{p}_i\,|A\rangle\Big|^2\frac{V\omega^2\,d\Omega}{(2\pi)^3\hbar c^3}, \tag{2.119}$$

where ω satisfies the energy conservation $E_A = E_B + \hbar\omega$. The normalization volume V cancels out as it should.

In a typical atomic transition in the optical region the wavelength of the emitted photon is much greater than the linear dimension of the atom:

$$\lambda\!\!\!^{-}_{\text{photon}} = 1/|\mathbf{k}| \gg r_{\text{atom}}, \tag{2.120}$$

since $\lambda\!\!\!^{-}_{\text{photon}}$ is typically of the order of several thousand angstrom units whereas the atomic radius is of the order of one angstrom unit. This means that we can replace

$$e^{-i\mathbf{k}\cdot\mathbf{x}_i} = 1 - i\mathbf{k}\cdot\mathbf{x}_i - (\mathbf{k}\cdot\mathbf{x}_i)^2/2 + \cdots \tag{2.121}$$

by its leading term 1. It turns out that the spin-magnetic-moment interaction is also negligible. To see this, just note that the matrix element of $(e/mc)\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{p}_i$ is larger than that of $(e\hbar/2mc)\boldsymbol{\sigma}_i\cdot(\mathbf{k}\times\boldsymbol{\epsilon}^{(\alpha)})$, again by $\lambda\!\!\!^{-}_{\text{photon}}/r_{\text{atom}}$ since the matrix element of $\mathbf{p}_i$ is of the order of $\hbar/r_{\text{atom}}$. An approximation in which only the $\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{p}_i$ term is kept is called the electric dipole ($E1$) approximation.

To further simplify the problem let us assume that only one of the atomic electrons participates in spontaneous emission, as in the case of a hydrogen-like atom (an atom in which there is only one valence electron).‡ Omitting the sum over i we have

$$w_{d\Omega} = \frac{e^2\omega}{8\pi^2m^2\hbar c^3}|\langle B|\,\mathbf{p}\,|A\rangle\cdot\boldsymbol{\epsilon}^{(\alpha)}|^2 d\Omega. \tag{2.122}$$

Meanwhile, using the commutation relation between $\mathbf{p}^2$ and $\mathbf{x}$,

$$[\mathbf{p}^2, \mathbf{x}] = -2i\hbar\mathbf{p}, \tag{2.123}$$

‡Many of the results we derive for one-electron atoms can readily be generalized to many-electron atoms.

we can rewrite the matrix element $\langle B \,|\, \mathbf{p} \,|\, A \rangle$ as follows:

$$\begin{aligned} \langle B \,|\, \mathbf{p} \,|\, A \rangle &= \langle B \left| \frac{im}{\hbar} [H_0, \mathbf{x}] \right| A \rangle \\ &= -\frac{im(E_B - E_A)}{\hbar} \langle B \,|\, \mathbf{x} \,|\, A \rangle \\ &= im\omega \mathbf{x}_{BA}. \end{aligned} \tag{2.124}$$

This matrix element is precisely what we would directly compute if the interaction Hamiltonian were replaced as follows:

$$-\frac{e\mathbf{A}\cdot\mathbf{p}}{mc} \longrightarrow \frac{e}{c}\mathbf{x}\cdot\frac{\partial \mathbf{A}}{\partial t} = -e\mathbf{x}\cdot\mathbf{E}, \tag{2.125}$$

provided that we again kept only the leading term in the plane-wave expansion of **E**. The origin of the term "electric dipole transition" is now evident.

The angular momentum selection rule for $E1$ is

$$|J_B - J_A| = 1, 0; \qquad \text{no} \quad 0 \longrightarrow 0. \tag{2.126}$$

To derive this we first note that the components of **x** can be rearranged as follows:

$$\begin{aligned} V^{\pm 1} &= \mp\frac{1}{\sqrt{2}}(x \pm iy) = r\sqrt{\frac{4\pi}{3}}\, Y_1^{\pm 1}, \\ V^0 &= z = r\sqrt{\frac{4\pi}{3}}\, Y_1^0. \end{aligned} \tag{2.127}$$

Then the Wigner-Eckart theorem gives

$$\langle J_B, m_B \,|\, V^q \,|\, J_A, m_A \rangle = \frac{1}{\sqrt{2J_A + 1}} \langle J_A 1 m_A q \,|\, J_B m_B \rangle \langle B \,||\, V^q \,||\, A \rangle, \tag{2.128}$$

where $\langle J_A 1 m_A q \,|\, J_B m_B \rangle$ is a Clebsch-Gordan coefficient.‡ The condition that the Clebsch-Gordan coefficient shall not vanish leads immediately to the angular momentum selection rule (2.126). Physically this selection rule is a consequence of the fact that one unit of angular momentum is carried away by the spin of the emitted photon in an $E1$ transition.

We also must consider the parity selection rule: parity change "yes." To prove this merely note that

$$\begin{aligned} \langle B \,|\, \mathbf{x} \,|\, A \rangle &= -\langle B \,|\, \Pi^{-1} \mathbf{x} \Pi \,|\, A \rangle \\ &= -\Pi_B \Pi_A \langle B \,|\, \mathbf{x} \,|\, A \rangle, \end{aligned} \tag{2.129}$$

where Π and $\Pi_{A,B}$ stand for the parity operator and its eigenvalue.

Equation (2.122) can also be written as

$$w_{d\Omega} = \frac{e^2\omega^3}{8\pi^2\hbar c^3} |\mathbf{x}_{BA}|^2 \cos^2 \Theta^{(\alpha)}\, d\Omega, \tag{2.130}$$

where we have defined the angle $\Theta^{(\alpha)}$ by

$$\cos \Theta^{(\alpha)} = |\mathbf{x}_{BA}\cdot\boldsymbol{\epsilon}^{(\alpha)}|/|\mathbf{x}_{BA}|. \tag{2.131}$$

---

‡Those who are unfamiliar with the Wigner-Eckart theorem and Clebsch-Gordan coefficients may consult Merzbacher (1961), Chapter 22; Messiah (1962), Chapter 13.

Note that

$$|\mathbf{x}_{BA}|^2 = |x_{BA}|^2 + |y_{BA}|^2 + |z_{BA}|^2$$
$$= \left|-\frac{(x_{BA} + iy_{BA})}{\sqrt{2}}\right|^2 + \left|\frac{x_{BA} - iy_{BA}}{\sqrt{2}}\right|^2 + |z_{BA}|^2. \tag{2.132}$$

So far we have been concerned with a radiative transition in which a photon with definite $\mathbf{k}$, α is emitted. We must now sum over two independent polarization states for given $\mathbf{k}$ and integrate over all propagation directions. From Fig. 2–1 it is evident that

$$\cos\Theta^{(1)} = \sin\theta\cos\phi, \qquad \cos\Theta^{(2)} = \sin\theta\sin\phi. \tag{2.133}$$

The sum over the two polarization states just gives $\sin^2\theta$. Integrating over all possible propagation directions with the direction of $\mathbf{x}_{BA}$ fixed in space, we obtain

$$2\pi\int_{-1}^{1}\sin^2\theta\, d(\cos\theta) = 8\pi/3. \tag{2.134}$$

Finally, for the integrated transition probability for spontaneous emission we have

$$w = \frac{e^2\omega^3}{3\pi\hbar c^3}|\mathbf{x}_{BA}|^2$$
$$= \left(\frac{e^2}{4\pi\hbar c}\right)\frac{4}{3}\frac{\omega^3}{c^2}|\mathbf{x}_{BA}|^2, \tag{2.135}$$

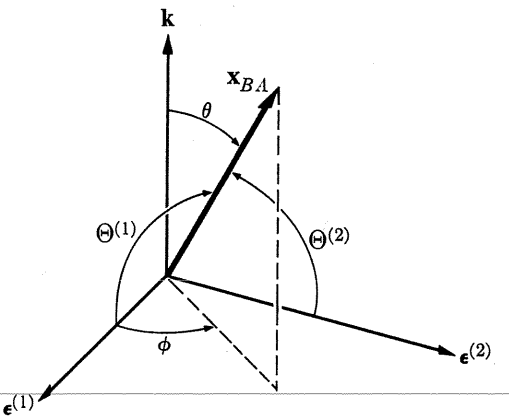

Fig. 2–1. The orientation of $\mathbf{x}_{BA}$.

with $(e^2/4\pi\hbar c) \simeq 1/137$. Note that the dimension of w is (time)$^{-1}$ as expected. This expression was obtained prior to the advent of quantum field theory by W. Heisenberg, who used the correspondence principle.

In order to compute the mean lifetime τ of state A we must sum the transition probabilities into all possible final states allowed by the selection rule and energy conservation:

$$1/\tau_A = \sum_i w_{A\to B_i}. \tag{2.136}$$

In particular, it is important to sum over the magnetic quantum number of the final state.

As a concrete example, we may mention that the mean lifetime of a radiative $E1$ transition from the state having the quantum numbers (n, l, m) to the state having (n', l', m') where $m' = m, m \pm 1$, is given by the reciprocal‡ of

$$\sum_{m'} w[(nlm)\longrightarrow(n'l'm')]$$
$$= \frac{e^2\omega^3}{3\pi\hbar c^3}\begin{Bmatrix}(l+1)/(2l+1)\\ l/(2l+1)\end{Bmatrix}\left|\int_0^\infty R_{n'l'}(r)R_{nl}(r)r^3\,dr\right|^2 \qquad \text{for} \qquad l' = \begin{cases}l+1\\ l-1\end{cases} \tag{2.137}$$

‡See Merzbacher (1961), p. 481, for a detailed derivation of this formula.

where $R_{nl}(r)$ is the normalized radial wave function of a hydrogen-like atom characterized by n and l. Note that the lifetime is independent of m. This result, which can be proved using the orthogonality properties of the Clebsch-Gordan coefficients (cf. Eq. 2.128), is reasonable in view of rotational invariance; the lifetime of an excited state in the absence of an external field should not depend on its orientation in space. In general, because the lifetime of an isolated state is independent of its m value, in computing the lifetime we can either sum over *just* the *final* magnetic quantum numbers for a fixed initial magnetic quantum number or sum over *both* the *initial and* the *final* magnetic quantum numbers and divide by the multiplicity of the initial state. For instance, in (2.137) we are free to make the replacement

$$\sum_{m'} \longrightarrow \frac{1}{2J+1} \sum_{m'} \sum_{m}, \tag{2.138}$$

where J (in this case, just l) is the angular momentum of the decaying state.

As an application of (2.137) we may compute the lifetime of the $2p$ state of the hydrogen atom which decays into the ground ($1s$) state. The radial wave functions are

$$R_{nl} = \frac{1}{\sqrt{24 a_0^3}} \frac{r}{a_0} e^{-r/2a_0} \quad \text{and} \quad R_{n'l'} = \frac{2}{\sqrt{a_0^3}} e^{-r/a_0}, \tag{2.139}$$

where a_0 is the Bohr radius. We obtain

$$\tau(2p \longrightarrow 1s) = 1.6 \times 10^{-9} \text{ sec.} \tag{2.140}$$

The mean lifetimes of other excited states of the hydrogen atom are tabulated in the literature.‡ In general, the lifetime of an excited hydrogen-like atom goes up with increasing n, roughly as n^3 for a fixed value of l, and as

$$\tau_n \propto \left(\sum_l \frac{(2l+1)}{n^5} \right)^{-1} \sim n^{4.5} \tag{2.141}$$

for the average over l.

Occasionally the symmetry of atomic states may be such that the electric dipole emission of a photon is forbidden. This occurs when $\mathbf{x}_{BA} = 0$ for every state B with an energy lower than A. It is then necessary to go back to the plane-wave expansion (2.121) and take seriously the term $\mathbf{k}\cdot\mathbf{x}$ which we previously ignored. The matrix element we must evaluate can be decomposed as follows:

$$\begin{aligned} \langle B | (\mathbf{k}\cdot\mathbf{x})(\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{p}) | A \rangle &= \tfrac{1}{2}\langle B | (\mathbf{k}\cdot\mathbf{x})(\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{p}) + (\mathbf{k}\cdot\mathbf{p})(\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{x}) | A \rangle \\ &\quad + \tfrac{1}{2}\langle B | (\mathbf{k}\cdot\mathbf{x})(\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{p}) - (\mathbf{k}\cdot\mathbf{p})(\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{x}) | A \rangle. \end{aligned} \tag{2.142}$$

The first term can be rewritten using

$$\tfrac{1}{2}[(\mathbf{k}\cdot\mathbf{x})(\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{p}) + (\mathbf{k}\cdot\mathbf{p})(\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{x})] = \tfrac{1}{2}\mathbf{k}\cdot(\mathbf{xp} + \mathbf{px})\cdot\boldsymbol{\epsilon}^{(\alpha)}, \tag{2.143}$$

where $\mathbf{xp} + \mathbf{px}$ is a symmetric dyadic. The radiative transition due to this term is known as an *electric quadrupole* ($E2$) transition since

$$\mathbf{xp} + \mathbf{px} = (im/\hbar)[H_0, \mathbf{xx}], \tag{2.144}$$

‡Bethe and Salpeter (1957), p. 266.

and
$$\frac{\mathbf{k}}{2}\cdot\langle B|\mathbf{xp}+\mathbf{px}|A\rangle\cdot\boldsymbol{\epsilon}^{(\alpha)}=-\frac{im\omega}{2}\mathbf{k}\cdot\langle B|\mathbf{xx}|A\rangle\cdot\boldsymbol{\epsilon}^{(\alpha)}. \tag{2.145}$$

Because of the transversality condition $\mathbf{k}\cdot\boldsymbol{\epsilon}^{(\alpha)}=0$, it is legitimate to replace $\mathbf{xx}$ by its traceless part whose individual components are

$$T_{ij}=x_ix_j-(\delta_{ij}/3)|\mathbf{x}|^2. \tag{2.146}$$

Note that it has five independent components which can be written as linear combinations of Y_2^m. It then follows from the Wigner-Eckhart theorem that the angular momentum selection rule for an $E2$ transition is

$$|J_B-J_A|\leq 2\leq J_B+J_A. \tag{2.147}$$

The second term in (2.142) can be written as follows:

$$(\mathbf{k}\cdot\mathbf{x})(\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{p})-(\mathbf{k}\cdot\mathbf{p})(\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{x})=(\mathbf{k}\times\boldsymbol{\epsilon}^{(\alpha)})\cdot(\mathbf{x}\times\mathbf{p}). \tag{2.148}$$

Now $\mathbf{k}\times\boldsymbol{\epsilon}^{(\alpha)}$ is the leading term in the plane-wave expansion of the magnetic field $\mathbf{B}$ while $\mathbf{x}\times\mathbf{p}$ is just the orbital angular momentum operator of the atomic electron which gives the orbital magnetic moment operator when combined with $e/(2mc)$. Hence the radiative transition due to this term is called a *magnetic dipole* ($M1$) *transition*. Together with this term we should consider the leading term from the spin magnetic moment interaction (2.95) $(e\hbar/2mc)\,\boldsymbol{\sigma}\cdot(\mathbf{k}\times\boldsymbol{\epsilon}^{(\alpha)})$, which is of the same order. The angular momentum selection rule for an $M1$ transition is $|J_B-J_A|\leq 1$, no $0\to 0$, just as in the $E1$ case. In contrast to an $E1$ transition, the parity of the atomic states changes neither in an $M1$ nor in an $E2$ transition.

We can treat higher multipole transitions by considering the higher powers of $\mathbf{k}\cdot\mathbf{x}$ in (2.121). However, it is better to use a formalism that employs *vector spherical harmonics*, a more powerful technique based on an expansion of $\boldsymbol{\epsilon}^{(\alpha)}e^{i\mathbf{k}\cdot\mathbf{x}}$ in terms of the eigenfunctions of the angular momentum operator of the radiation field. We shall not discuss this method; it is treated in standard textbooks on nuclear physics.‡

The atomic states for which electric dipole transitions are forbidden have long lifetimes. In contrast to the lifetime of order 10^{-8} sec characteristic of a typical $E1$ transition, the lifetimes of typical $M1$ or $E2$ transitions are of the order of 10^{-3} sec, roughly $(\lambda\!\!\!^{-}/r_{\text{atom}})^2$ times the typical $E1$ lifetime, as expected. As an example in which $E1$, $M1$, $E2$, etc., are all forbidden for practical purposes, we may mention the radiative transition between the metastable $2s$ state and the ground ($1s$) state of the hydrogen atom. For this transition $E1$ is forbidden by parity, the $M1$ matrix element vanishes when the nonrelativistic wave functions are used, and $E2$ and all other higher multipole transitions are forbidden by angular momentum conservation. Its decay mode turns out to be the simultaneous emission of two photons which can be calculated to be $\frac{1}{7}$ sec (cf. Problem 2–6). Note that this lifetime is extremely long compared to the lifetime of the $2p$ state, 1.6×10^{-9} sec.

‡Consult, for example, Blatt and Weisskopf (1952), Chapter 12.

Planck's radiation law. We wish to conclude this section with a derivation of Planck's radiation law from the point of view of quantum field theory. Suppose we have atoms and a radiation field which can freely exchange energy by the reversible process

$$\mathrm{A} \rightleftarrows \gamma + \mathrm{B} \tag{2.149}$$

in such a way that thermal equilibrium is established. If the populations of the upper and lower atomic levels are denoted by $N(A)$ and $N(B)$ respectively, we have the equilibrium condition

$$N(B)w_{\text{abs}} = N(A)w_{\text{emis}}, \tag{2.150}$$

$$\frac{N(B)}{N(A)} = \frac{e^{-E_B/kT}}{e^{-E_A/kT}} = e^{\hbar\omega/kT}, \tag{2.151}$$

where w_{abs} and w_{emis} are respectively the transition probabilities for $B + \gamma \to A$ and $A \to B + \gamma$. According to (2.96), (2.99), and (2.114) we have

$$\frac{w_{\text{emis}}}{w_{\text{abs}}} = \frac{(n_{\mathbf{k},\alpha} + 1)\,|\sum_i \langle B|\, e^{-i\mathbf{k}\cdot\mathbf{x}_i}\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{p}_i\,|A\rangle|^2}{n_{\mathbf{k},\alpha}\,|\sum_i \langle A|\, e^{i\mathbf{k}\cdot\mathbf{x}_i}\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{p}_i\,|B\rangle|^2}. \tag{2.152}$$

But

$$\langle B|\, e^{-i\mathbf{k}\cdot\mathbf{x}_i}\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{p}_i\,|A\rangle = \langle A|\, \mathbf{p}_i\cdot\boldsymbol{\epsilon}^{(\alpha)} e^{i\mathbf{k}\cdot\mathbf{x}_i}|B\rangle^* = \langle A|\, e^{i\mathbf{k}\cdot\mathbf{x}_i}\boldsymbol{\epsilon}^{(\alpha)}\cdot\mathbf{p}_i\,|B\rangle^*. \tag{2.153}$$

Hence

$$\frac{N(B)}{N(A)} = \frac{w_{\text{emis}}}{w_{\text{abs}}} = \frac{n_{\mathbf{k},\alpha} + 1}{n_{\mathbf{k},\alpha}}. \tag{2.154}$$

Using (2.151) and (2.154) we obtain

$$n_{\mathbf{k},\alpha} = \frac{1}{e^{\hbar\omega/kT} - 1}. \tag{2.155}$$

All this is for photon states which satisfy $\hbar\omega = E_A - E_B$. Suppose the radiation field is enclosed by "black" walls which are made up of various kinds of atoms and are capable of absorbing and re-emitting photons of any energy. The energy of the radiation field, per volume, in the angular frequency interval $(\omega, \omega + d\omega)$ is given by (cf. Eq. 2.116)

$$\begin{aligned} U(\omega)\,d\omega &= \frac{1}{L^3}\left(\frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}\right) 2\left(\frac{L}{2\pi}\right)^3 4\pi k^2\,dk \\ &= \frac{8\pi\hbar}{c^3}\left(\frac{\omega}{2\pi}\right)^3\left(\frac{1}{e^{\hbar\omega/kT} - 1}\right) d\omega. \end{aligned} \tag{2.156}$$

The energy distribution per frequency per volume is

$$U(\nu) = U(\omega)\frac{d\omega}{d\nu} = \frac{8\pi h\nu^3}{c^3}\frac{1}{e^{h\nu/kT} - 1}. \tag{2.157}$$

This is Planck's famous law which opened up the twentieth-century physics of the quantum domain.

It is instructive to compare our derivation of Planck's law with Einstein's 1917 derivation.‡ They are both based on thermal equilibrium between the atoms and the radiation field. In Einstein's derivation the principle of detailed balance is *explicitly* invoked; by contrast, in our derivation the physics of detailed balance is contained in (2.153) which is an automatic consequence of the hermiticity of the Hamiltonian used in the quantum theory of radiation. Note also that in our derivation we do not distinguish between the contributions from spontaneous emission and induced emission.

Although our attention has been focused in this section on the radiative transitions between two atomic states, the techniques we have acquired can readily be applied to a host of other phenomena. For instance, the reader may calculate the cross section for the photoelectric effect (Problem 2–4) or the lifetime of the Σ^0 hyperon (Problem 2–5):

$$\Sigma^0 \overset{M1}{\rightarrow} \Lambda + \gamma.$$

2–5. RAYLEIGH SCATTERING, THOMSON SCATTERING, AND THE RAMAN EFFECT

Kramers-Heisenberg formula. Let us now examine the field-theoretic treatment of the scattering of photons by atomic electrons. Before the scattering, the atom is in state A, and the incident photon is characterized by $(\mathbf{k}, \boldsymbol{\epsilon}^{(\alpha)})$. After the scattering, the atom is left in state B, and the outgoing photon is characterized by $(\mathbf{k}', \boldsymbol{\epsilon}^{(\alpha')})$. For simplicity let us again consider a one-electron atom and neglect the spin-magnetic-moment interaction.

The interaction Hamiltonian (2.94) is made up of a linear $(\mathbf{A}\cdot\mathbf{p})$ term and a quadratic $(\mathbf{A}\cdot\mathbf{A})$ term. Since $\mathbf{A}$ changes the number of photons by one, $\mathbf{A}\cdot\mathbf{p}$ makes no contribution in first order to a scattering process in which there is no *net* change in the number of photons. On the other hand, the $\mathbf{A}\cdot\mathbf{A}$ term contains $aa^\dagger$, $a^\dagger a$, $a\,a$, and $a^\dagger a^\dagger$, the first two of which *do* give nonvanishing contributions provided that $a^\dagger$ and a, respectively, represent the creation operator for $(\mathbf{k}', \alpha')$ and the annihilation operator for $(\mathbf{k}, \alpha)$, e.g., $\langle \mathbf{k}', \alpha' \,|\, a_{\mathbf{k},\alpha}\, a^\dagger_{\mathbf{k}',\alpha'} \,|\, \mathbf{k}, \alpha\rangle = 1$. Hence

$$\begin{aligned}
&\langle B; \mathbf{k}', \boldsymbol{\epsilon}^{(\alpha')} \,|\, H_{\text{int}} \,|\, A; \mathbf{k}, \boldsymbol{\epsilon}^{(\alpha)}\rangle \\
&\quad = \langle B; \mathbf{k}', \boldsymbol{\epsilon}^{(\alpha')} \left| \frac{e^2}{2mc^2} \mathbf{A}(\mathbf{x}, t)\cdot\mathbf{A}(\mathbf{x}, t) \right| A; \mathbf{k}, \boldsymbol{\epsilon}^{(\alpha)}\rangle \\
&\quad = \langle B; \mathbf{k}', \boldsymbol{\epsilon}^{(\alpha')} \left| \frac{e^2}{2mc^2} (a_{\mathbf{k},\alpha} a^\dagger_{\mathbf{k}',\alpha'} + a^\dagger_{\mathbf{k}',\alpha'} a_{\mathbf{k},\alpha}) \frac{c^2\hbar}{2V\sqrt{\omega\omega'}} \boldsymbol{\epsilon}^{(\alpha)}\cdot\boldsymbol{\epsilon}^{(\alpha')} \right. \\
&\qquad\qquad \times \exp\left[i(\mathbf{k} - \mathbf{k}')\cdot\mathbf{x} - i(\omega - \omega')t\right] \,|\, A; \mathbf{k}, \boldsymbol{\epsilon}^{(\alpha)}\rangle \\
&\quad = \frac{e^2}{2mc^2} \frac{c^2\hbar}{2V\sqrt{\omega\omega'}} 2\boldsymbol{\epsilon}^{(\alpha)}\cdot\boldsymbol{\epsilon}^{(\alpha')} \exp\left[-i(\omega - \omega')t\right] \langle B \,|\, A\rangle, \qquad (2.158)
\end{aligned}$$

‡See, for example, Kittel (1958), pp. 175–176.

where we have replaced $e^{i\mathbf{k}\cdot\mathbf{x}}$ and $e^{-i\mathbf{k}'\cdot\mathbf{x}}$ by 1, since in the long-wave approximation the atomic electron may be assumed to be situated at the origin. For the first-order transition amplitude $c^{(1)}(t)$ we have

$$c^{(1)}(t) = \frac{1}{i\hbar}\frac{e^2}{2mc^2}\frac{c^2\hbar}{2V\sqrt{\omega\omega'}}2\delta_{AB}\boldsymbol{\epsilon}^{(\alpha)}\cdot\boldsymbol{\epsilon}^{(\alpha')}\int_0^t \exp\left[i(\hbar\omega' + E_B - \hbar\omega - E_A)t_1/\hbar\right]dt_1, \tag{2.159}$$

with $\omega = |\mathbf{k}|c$ and $\omega' = |\mathbf{k}'|c$ as usual.

Although the $\mathbf{A}\cdot\mathbf{p}$ term makes no contribution in first order, the $\mathbf{A}\cdot\mathbf{p}$ term taken *twice* is of the same order as the $\mathbf{A}\cdot\mathbf{A}$ term, so far as powers of e are concerned. Therefore we must treat a double $\mathbf{A}\cdot\mathbf{p}$ interaction and a single $\mathbf{A}\cdot\mathbf{A}$ interaction simultaneously. The $\mathbf{A}\cdot\mathbf{p}$ interaction acting at t_1 can either annihilate the incident photon $(\mathbf{k}, \alpha)$ or create the outgoing photon $(\mathbf{k}', \alpha')$. When the $\mathbf{A}\cdot\mathbf{p}$ interaction acts again at a time t_2 which is later than t_1 it must necessarily create the outgoing photon $(\mathbf{k}', \alpha')$ if the outgoing photon has not yet been created. Otherwise we would end up with a zero matrix element. On the other hand, if the outgoing photon has already been created but the incoming photon has not yet been annihilated, the $\mathbf{A}\cdot\mathbf{p}$ interaction acting at $t_2 > t_1$ must annihilate the incoming photon $(\mathbf{k}, \alpha)$. Between t_1 and t_2 the atom is in state I which is, in general, different from A and B. To summarize, two types of intermediate states are possible. In the first type the atom is in state I and no photons are present. In the second type the atom is in state I and both the incident and the outgoing photon are present.‡

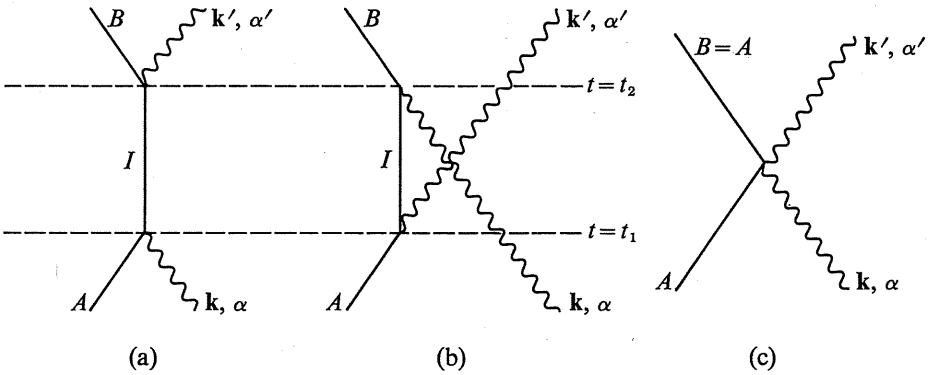

Fig. 2–2. Space-time diagram for scattering of light.

All this can best be visualized if we draw a space-time diagram (Feynman diagram) in which a solid line represents the atom, and a wavy line represents a photon. Time is assumed to run upward (Fig. 2–2). For a type 1 process, represented by Fig. 2–2(a), the atomic state A first absorbs the incident photon at t_1 and becomes state I; subsequently at t_2 the atomic state I emits the outgoing photon and changes

‡Strictly speaking, we should also consider the case where I stands for a continuum state. The relevant matrix element then corresponds to a photo-effect matrix element (cf. Problem 2–4). In practice such "distant" intermediate states are not important because the energy denominators become large (cf. Eq. 2.160 below).

into state B. For a type 2 process, represented by Fig. 2–2(b), state A first emits the outgoing photon at t_1 and changes into state I; subsequently at t_2 state I absorbs the incident photon (which has not yet been annihilated) and becomes state B. In contrast, the lowest-order $\mathbf{A}\cdot\mathbf{A}$ interaction, discussed earlier, is represented by Fig. 2–2(c) ("seagull graph").

As emphasized in the previous section, the emission and absorption of a photon by an atomic electron are equivalent to interactions of the atomic electron with the time-dependent potentials (2.102). Using this rule, we can readily write down the second-order transition amplitude $c^{(2)}(t)$ as follows:

$$\begin{aligned} c^{(2)}(t) = \frac{1}{(i\hbar)^2}\frac{c^2\hbar}{2V\sqrt{\omega\omega'}}\left(-\frac{e}{mc}\right)^2 \int_0^t dt_2 \int_0^{t_2} dt_1 \\ \times [\sum_I \langle B|\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')}|I\rangle \exp[i(E_B - E_I + \hbar\omega')t_2/\hbar] \\ \times \langle I|\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)}|A\rangle \exp[i(E_I - E_A - \hbar\omega)t_1/\hbar] \\ + \sum_I \langle B|\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)}|I\rangle \exp[i(E_B - E_I - \hbar\omega)t_2/\hbar] \\ \times \langle I|\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')}|A\rangle \exp[i(E_I - E_A + \hbar\omega')t_1/\hbar] \\ = -\frac{c^2\hbar}{i\hbar 2V\sqrt{\omega\omega'}}\left(\frac{e}{mc}\right)^2 \sum_I \left(\frac{(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{BI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{IA}}{E_I - E_A - \hbar\omega} + \frac{(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{BI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{IA}}{E_I - E_A + \hbar\omega'}\right) \\ \times \int_0^t dt_2 \exp[i(E_B - E_A + \hbar\omega' - \hbar\omega)t_2/\hbar], \end{aligned} \tag{2.160}$$

where we have made the dipole approximation and ignored a term that depends on the artificial sudden turning of the perturbation (which is negligible if the energy conservation, $E_B - E_A + \hbar\omega' - \hbar\omega = 0$, is nearly satisfied). Combining $c^{(1)}(t)$ and $c^{(2)}(t)$, we have the transition probability

$$\begin{aligned} w_{d\Omega} &= \int (|c^{(1)} + c^{(2)}|^2/t)\rho_{E,d\Omega}\, dE \\ &= \frac{2\pi}{\hbar}\left(\frac{c^2\hbar}{2V\sqrt{\omega\omega'}}\right)^2\left(\frac{e^2}{mc^2}\right)^2 \frac{V}{(2\pi)^3\hbar c^3}\omega'^2\, d\Omega \\ \times \left|\delta_{AB}\boldsymbol{\epsilon}^{(\alpha)}\cdot\boldsymbol{\epsilon}^{(\alpha')} - \frac{1}{m}\sum_I \left(\frac{(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{BI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{IA}}{E_I - E_A - \hbar\omega} + \frac{(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{BI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{IA}}{E_I - E_A + \hbar\omega'}\right)\right|^2 \cdot \end{aligned} \tag{2.161}$$

To obtain the differential cross section we must divide this transition probability by the flux density which is just c/V, since initially there is one photon in the normalization box of volume V. Finally, we have for the differential cross section

$$\begin{aligned} \frac{d\sigma}{d\Omega} = r_0^2\left(\frac{\omega'}{\omega}\right)\left|\delta_{AB}\boldsymbol{\epsilon}^{(\alpha)}\cdot\boldsymbol{\epsilon}^{(\alpha')} - \frac{1}{m}\sum_I \left(\frac{(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{BI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{IA}}{E_I - E_A - \hbar\omega}\right.\right. \\ \left.\left. + \frac{(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{BI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{IA}}{E_I - E_A + \hbar\omega'}\right)\right|^2, \end{aligned} \tag{2.162}$$

where r_0 stands for the classical radius of the electron, and

$$r_0 = \frac{e^2}{4\pi mc^2} \simeq \frac{1}{137}\frac{\hbar}{mc} \simeq 2.82 \times 10^{-13}\ \text{cm}. \tag{2.163}$$

A formula equivalent to (2.162) was first obtained by H. A. Kramers and W. Heisenberg using the correspondence principle in 1925; hence it is called the Kramers-Heisenberg formula.

Rayleigh scattering. There are certain special cases of (2.162) worth examining in detail. Let us first discuss the case in which $A = B$, $\hbar\omega = \hbar\omega'$. This situation corresponds to elastic scattering of light. It is also called Rayleigh scattering because this problem was treated classically by Lord Rayleigh. To simplify (2.162) we rewrite $\boldsymbol{\epsilon}^{(\alpha)}\cdot\boldsymbol{\epsilon}^{(\alpha')}$, using the commutation relation between $\mathbf{x}$ and $\mathbf{p}$, the completeness of the intermediate states I, and (2.124):‡

$$\begin{aligned}\boldsymbol{\epsilon}^{(\alpha)}\cdot\boldsymbol{\epsilon}^{(\alpha')} &= \frac{1}{i\hbar}\sum_I [(\mathbf{x}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{AI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{IA} - (\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{AI}(\mathbf{x}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{IA}] \\ &= \frac{1}{m\hbar}\sum_I \frac{1}{\omega_{IA}}[(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{AI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{IA} + (\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{AI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{IA}] \qquad (2.164)\end{aligned}$$

where $\omega_{IA} = (E_I - E_A)/\hbar$. We now see that the three terms in (2.162) combine so that

$$\begin{aligned}&\delta_{AA}\boldsymbol{\epsilon}^{(\alpha)}\cdot\boldsymbol{\epsilon}^{(\alpha')} - \frac{1}{m\hbar}\sum_I \left[\frac{(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{AI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{IA}}{\omega_{IA}-\omega} + \frac{(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{AI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{IA}}{\omega_{IA}+\omega}\right] \\ &\quad = -\frac{1}{m\hbar}\sum_I \left[\frac{\omega(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{AI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{IA}}{\omega_{IA}(\omega_{IA}-\omega)} - \frac{\omega(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{AI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{IA}}{\omega_{IA}(\omega_{IA}+\omega)}\right]. \qquad (2.165)\end{aligned}$$

Using the expansion $1/(\omega_{IA} \mp \omega) \approx [1 \pm (\omega/\omega_{IA})]/\omega_{IA}$, valid for small values of ω, and

$$\begin{aligned}&\sum_I \frac{1}{\omega_{IA}^2}[(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{AI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{IA} - (\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{AI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{IA}] \\ &\quad = m^2\sum_I [(\mathbf{x}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{AI}(\mathbf{x}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{IA} - (\mathbf{x}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{AI}(\mathbf{x}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{IA}] \\ &\quad = m^2([\mathbf{x}\cdot\boldsymbol{\epsilon}^{(\alpha')}, \mathbf{x}\cdot\boldsymbol{\epsilon}^{(\alpha)}])_{AA} \\ &\quad = 0, \qquad (2.166)\end{aligned}$$

we obtain the Rayleigh cross section for $\omega \ll \omega_{IA}$:

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= \left(\frac{r_0}{m\hbar}\right)^2\omega^4\left|\sum_I\left(\frac{1}{\omega_{IA}}\right)^3[(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{AI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{IA} + (\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{AI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})]_{IA}\right|^2 \\ &= \left(\frac{r_0 m}{\hbar}\right)^2\omega^4\left|\sum_I\left(\frac{1}{\omega_{IA}}\right)[(\mathbf{x}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{AI}(\mathbf{x}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{IA} + (\mathbf{x}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{AI}(\mathbf{x}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{IA}]\right|^2. \qquad (2.167)\end{aligned}$$

Thus we see that *the scattering cross section at long wavelengths varies as the inverse fourth power of the wavelength* (Rayleigh's law). For atoms in ordinary colorless gases the light wave corresponding to a typical ω_{IA} is in the ultraviolet region.

‡The intermediate states I form a complete set only when we include the continuum states as well as the discrete (bound) states.

Hence the approximation $\omega \ll \omega_{IA}$ is good for ω in the visible optical region. This theory explains why the sky is blue and the sunset is red.

Thomson scattering. Let us now consider the opposite case in which the incident photon energy is much larger than the atomic binding energy. It is then legitimate to ignore the second and third term of (2.162), since $\hbar\omega(=\hbar\omega')$ is much larger than $(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{AI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{IA}/m$, so the scattering is due solely to the matrix element corresponding to the "seagull graph" (Fig. 2–2c). Now the $\delta_{AB}\boldsymbol{\epsilon}^{(\alpha)}\cdot\boldsymbol{\epsilon}^{(\alpha')}$ term is insensitive to the nature of the binding of the atomic electron. The cross section we compute in this case coincides with the cross section for the scattering of light by a free (unbound) electron, first obtained classically by J. J. Thomson:

$$\frac{d\sigma}{d\Omega} = r_0^2\,|\boldsymbol{\epsilon}^{(\alpha)}\cdot\boldsymbol{\epsilon}^{(\alpha')}|^2. \tag{2.168}$$

Note that this expression is independent of ω.

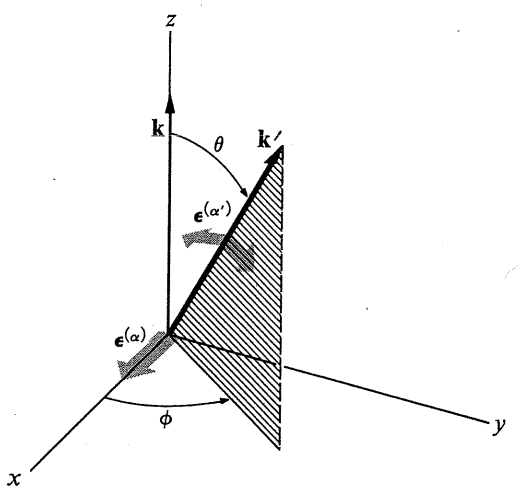

Fig. 2–3. Polarization in Thomson scattering.

To study the polarization dependence of Thomson scattering we consider a coordinate system in which $\boldsymbol{\epsilon}^{(\alpha)}$ and $\mathbf{k}$ are taken along the x-and the z-axes respectively, as shown in Fig. 2–3. The orientation of $\mathbf{k}'$ is characterized by the spherical coordinate angles θ and ϕ. The final polarization vector $\boldsymbol{\epsilon}^{(\alpha')}$ may be taken to be *normal* to the shaded plane (the plane determined by $\mathbf{k}$ and $\mathbf{k}'$) for $\alpha' = 1$; $\boldsymbol{\epsilon}^{(\alpha')}$ with $\alpha' = 2$ must then lie *in* the shaded plane. The Cartesian components of $\boldsymbol{\epsilon}^{(\alpha')}$ are given by

$$\boldsymbol{\epsilon}^{(\alpha')} = \begin{cases} (\sin\phi, -\cos\phi, 0) & \text{for } \alpha' = 1, \\ (\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta) & \text{for } \alpha' = 2. \end{cases} \tag{2.169}$$

Hence

$$\frac{d\sigma}{d\Omega} = r_0^2 \begin{cases} \sin^2\phi & \text{for } \alpha' = 1, \\ \cos^2\theta\cos^2\phi & \text{for } \alpha' = 2. \end{cases} \tag{2.170}$$

For initially unpolarized photons we may either integrate (2.170) over the angle ϕ and divide by 2π or evaluate

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{unpolarized}} = \frac{1}{2}\left[\frac{d\sigma}{d\Omega}(\phi = 0) + \frac{d\sigma}{d\Omega}\left(\phi = \frac{\pi}{2}\right)\right]. \tag{2.171}$$

The two procedures are completely equivalent. Note that even if the initial polarization vector is randomly oriented, the final photon emitted with $\cos\theta \neq \pm 1$ is polarized, since the differential cross section is $r_0^2/2$ for $\boldsymbol{\epsilon}^{(\alpha')}$ normal to the plane determined by $\mathbf{k}$ and $\mathbf{k}'$ and $(r_0^2/2)\cos^2\theta$ for $\boldsymbol{\epsilon}^{(\alpha')}$ lying in the plane. It is remarkable that the polarization of the scattered photon is complete for $\theta = \pi/2$. We find, then, that a completely unpolarized light beam, when scattered through 90°, results in a 100 % linearly polarized beam whose polarization vector is normal to the plane determined by $\mathbf{k}$ and $\mathbf{k}'$.

If the initial photon is polarized but the final photon polarization is not observed, we must sum over the two possible states of polarization. We have

$$\frac{d\sigma}{d\Omega}\bigg|_{\substack{\text{final}\\ \text{polarization}\\ \text{summed}}} = r_0^2(\sin^2\phi + \cos^2\theta\cos^2\phi). \tag{2.172}$$

If the initial photon is not polarized and the final photon polarization is not measured, we have

$$\frac{d\sigma}{d\Omega}\bigg|_{\substack{\text{unpolarized;}\\ \text{final polarization}\\ \text{summed}}} = \frac{r_0^2}{2}(1 + \cos^2\theta). \tag{2.173}$$

The total cross section for Thomson scattering is

$$\sigma_{\text{tot}} = \frac{8\pi r_0^2}{3} = 6.65 \times 10^{-25}\ \text{cm}^2. \tag{2.174}$$

As we emphasized earlier, this expression for the cross section is valid at photon energies much greater than the atomic binding energy. However, the foregoing derivation breaks down if the photon energy is so high that it actually becomes comparable to the rest energy of the electron. We must then take into account the relativistic nature of the electron, as we shall do in Section 4–4, discussing Compton scattering.

The quantum-theoretic treatment of Rayleigh and Thomson scattering can be compared to the classical counterpart. The scattering of an electromagnetic wave can be visualized in classical mechanics by the following two-step process:

a) A bound electron oscillates when it is exposed to a time-dependent electric field.
b) The oscillating charge in turn radiates an electromagnetic wave.

For a model of the electron bound by a force obeying Hooke's law, the displacement $\mathbf{x}$ of the electron in the presence of an applied electric field $\mathbf{E}_0 e^{-i\omega t}$ satisfies the differential equation

$$\ddot{\mathbf{x}} + \omega_0^2\mathbf{x} = (e/m)\mathbf{E}_0 e^{-i\omega t}, \tag{2.175}$$

where ω_0 is the characteristic angular frequency of the oscillator. Knowing that the acceleration of the electron is given by

$$\ddot{\mathbf{x}} = -\left(\frac{e}{m}\right)\left(\frac{\omega^2}{\omega_0^2 - \omega^2}\right)\mathbf{E}_0 e^{-i\omega t}, \tag{2.176}$$

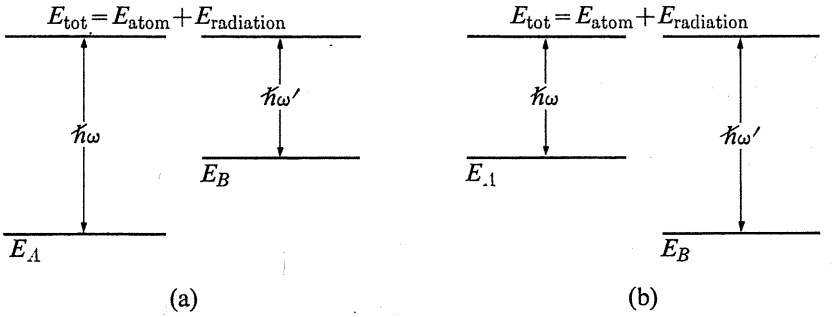

Fig. 2–4. (a) Stokes' line, (b) anti-Stokes' line.

we can readily compute the total scattering cross section in a straightforward manner.‡ We obtain

$$\sigma_{tot} = \frac{8\pi r_0^2}{3} \frac{\omega^4}{(\omega_0^2 - \omega^2)^2}. \tag{2.177}$$

For $\omega \ll \omega_0$ we have the ω^4 dependence of (2.167), whereas for $\omega \gg \omega_0$ we recover the frequency independent cross section (2.174).

The Raman effect. The Kramers-Heisenberg formula (2.162) can also be applied to inelastic scattering of light in which $\omega \neq \omega'$ and $A \neq B$. In atomic physics this phenomenon is called the *Raman effect* after C. V. Raman who observed a shift in the frequency of radiation scattered in liquid solutions, an effect predicted earlier by A. Smekal. If the initial atomic state A is the ground state, then the energy of the final photon $\hbar\omega'$ cannot be greater than the incident photon energy $\hbar\omega$ because $\hbar\omega + E_A = \hbar\omega' + E_B$ (Fig. 2–4a). This accounts for the presence of a *Stokes' line* in atomic spectra, a spectral line more reddish than that of incident radiation. On the other hand, if the atom is in an excited state, ω' can be larger than ω (Fig. 2–4b). This leads to an *anti-Stokes' line* which is more violet than the spectral line of the incident radiation.

2–6. RADIATION DAMPING AND RESONANCE FLUORESCENCE

The Kramers-Heisenberg formula we derived in the previous section is clearly inadequate if $\hbar\omega$ becomes equal to $E_I - E_A$ for some state I. The cross section according to (2.162) is then infinite, a phenomenon not observable in nature, of course. It is nevertheless true that the scattering cross section becomes very large and goes through a very sharp maximum in the neighborhood of $E_I - E_A = \hbar\omega$. This is a phenomenon known as resonance scattering of light or *resonance fluorescence.*

Where did our theory go wrong? When we use the second-order time-dependent perturbation theory, we assumed that the intermediate state I is a stationary state with an infinitely long lifetime. In other words, we did not take into account the

‡Panofsky and Phillips (1955), p. 326; Jackson (1962), pp. 602–604.

instability of state I due to spontaneous emission. In this connection, it is amusing to note that the classical expression for the Rayleigh scattering cross section (2.177) also blows up for $\omega \approx \omega_0$. This catastrophe can be avoided by introducing a damping force in the differential equation (2.175) as follows:

$$\ddot{\mathbf{x}} + \gamma\dot{\mathbf{x}} + \omega_0^2\mathbf{x} = (e/m)\mathbf{E}_0 e^{-i\omega t}. \tag{2.178}$$

The scattering cross section is then given by

$$\sigma_{\text{tot}} = \frac{8\pi r_0^2}{3}\frac{\omega^4}{(\omega_0^2 - \omega^2)^2 + \gamma^2/4}, \tag{2.179}$$

which is large but finite at $\omega = \omega_0$. As we shall see in a moment, in the quantum-theoretic treatment the depletion of state I due to spontaneous emission plays a role analogous to the damping force in the classical treatment; hence it is known as *radiation damping*. The quantum theory of unstable states was first developed by V. F. Weisskopf and E. P. Wigner precisely in this connection.

To simplify, let us assume that the initial atomic state A is a stable (ground) state. As the light beam hits the atom, the coefficient $c_I(t)$ builds up because state A makes a transition to state I as it absorbs the incident photon. Mathematically the differential equation for $c_I(t)$ due to this absorption process is (cf. Eq. 2.107)

$$\dot{c}_I = \frac{1}{i\hbar} H_{IA}^{(\text{abs})}(t) c_A e^{i(E_I - E_A)/\hbar}, \tag{2.180}$$

where $H_{IA}^{\text{abs}}(t)$ is the time-dependent matrix element characteristic of the photon absorption. But this is not the whole story. The amplitude $c_I(t)$ changes with time because of spontaneous emission even if state I is left alone in the absence of any incident radiation. The probability of finding the excited state I decreases with time as $e^{-\Gamma_I t/\hbar}$, where $\Gamma_I = \hbar/\tau_I$, τ_I being the mean life of state I given by (2.135) and (2.136). Since the amplitude for state I must vary as $e^{-\Gamma_I t/2\hbar}$ in the absence of any incident radiation, we can take into account the depletion of state I by adding to the right-hand side of the differential equation (2.180) a new term as follows:

$$\dot{c}_I = \frac{1}{i\hbar} H_{IA}^{(\text{abs})}(t) c_A e^{i(E_I - E_A)t/\hbar} - \frac{\Gamma_I}{2\hbar} c_I. \tag{2.181}$$

The first term of (2.181) characterizes the growth of state I due to the absorption of light, while the second term describes the depletion of state I due to spontaneous emission. Admittedly our approach is somewhat phenomenological; a more rigorous justification of the second term will be given in Section 2–8 when we discuss the second-order level shift due to the emission and absorption of a virtual photon.

We solve the differential equation (2.181) subject to the initial condition $c_I(0) = 0$, $c_A(0) = 1$. By direct substitution we find that the solution to this differential equation satisfying the above initial condition is

$$c_I^{(1)}(t) = \frac{H'_{IA}\left(\exp\left[-\Gamma_I t/2\hbar\right] - \exp\left[i(E_I - E_A - \hbar\omega)t/\hbar\right]\right)}{E_I - E_A - \hbar\omega - i\Gamma_I/2}, \tag{2.182}$$

where H'_{IA} is the time-independent matrix element given by

$$H_{IA}^{(\text{abs})}(t) = H'_{AI} e^{-i\omega t} = -(e/mc)\langle I \mid c\sqrt{\hbar/2\omega V}\,\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)} \mid A\rangle e^{-i\omega t}. \tag{2.183}$$

The lowest-order probability for finding state I is just $|c_I^{(1)}(t)|^2$. Note that, unlike (2.114) of Section 2–4, it is *not* proportional to t times $\delta(E_I - E_A - \hbar\omega)$ as $t \to \infty$. Whether (2.182) or (2.113) is the more realistic expression for describing a given absorption process depends on whether the lifetime of state I is short or long compared to the observation time. From (2.182) we obtain

$$|c_I^{(1)}(\infty)|^2 = \frac{|H'_{IA}|^2}{(E_I - E_A - \hbar\omega)^2 + \Gamma_I^2/4}. \tag{2.184}$$

State I is found copiously in abundance when the resonance condition $E_I - E_A \approx \hbar\omega$ is satisfied. If (2.184) is regarded as a function of the incident photon energy $\hbar\omega$, it has the familiar Lorentz shape whose full width at half maximum is given by $\Gamma_I = \hbar/\tau_I$. When the atom is irradiated by a continuous spectrum of radiation, we expect an absorption line whose width corresponds in the energy scale to Γ_I, often called the *natural width*. Using a similar argument, we can show that the frequency dependence of an emission line for an excited state I returning to the ground state A is also given by the same form, as we would expect from considerations based on equilibrium (Kirchhoff's law).‡

Let us now return to the scattering problem. To compute the scattering cross section for $\gamma + A \to \gamma + B$, where $\hbar\omega \simeq E_I - E_A$ for some I, we must re-evaluate $c_B^{(2)}$ using our modified $c_I^{(1)}$ in the general formula (cf. Eq. 2.111).

$$c_B^{(2)}(t) = \frac{1}{i\hbar}\sum_I \int_0^t dt'' \langle B|H_I(t'')|I\rangle\, e^{i(E_B - E_I)t''/\hbar} c_I^{(1)}(t''). \tag{2.185}$$

For a sufficiently large value of t we can omit the transient term $e^{-\Gamma_I t/2\hbar}$ in (2.182). Denoting by R those intermediate states I for which the resonance condition $\hbar\omega \simeq E_I - E_A$ is satisfied, we obtain the second-order amplitude $c_B^{(2)}$ as follows:

$$c_B^{(2)}(t) = -\frac{1}{i\hbar}\left(-\frac{e}{mc}\right)^2 \sum_R \frac{\langle B|c\sqrt{\hbar/2\omega' V}\,\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')}|R\rangle\langle R|c\sqrt{\hbar/2\omega V}\,\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)}|A\rangle}{E_R - E_A - \hbar\omega - i\Gamma_R/2}$$

$$\times \int_0^t dt'' \exp[i(E_B - E_A + \hbar\omega' - \hbar\omega)t''/\hbar] + \text{nonresonant terms}, \tag{2.186}$$

where the "nonresonant terms" stand for (2.160) except that in the sum over I for the type 1 intermediate states those states (R) which satisfy the resonance condition are missing. We see that the only change necessary is the substitution

$$E_I \longrightarrow E_I - i(\Gamma_I/2) \tag{2.187}$$

---

‡The foregoing discussion of line width ignores the broadenings of spectral lines due to the Doppler effect and atomic collisions which, in many instances, are more important than the broadening due to spontaneous emission. However, our phenomenological treatment can easily be generalized to the case where the depletion of state I is due to inelastic collisions. Compare our treatment with the discussion on absorption of radiation found in Dicke and Wittke (1960), pp. 273–275, where the main depletion is assumed to be due to atomic collisions rather than spontaneous emission.

for the type 1 intermediate states satisfying the resonance condition. Hence for any arbitrary ω we have the modified Kramers-Heisenberg formula

$$\frac{d\sigma}{d\Omega} = r_0^2\left(\frac{\omega'}{\omega}\right)\left|\delta_{AB}\boldsymbol{\epsilon}^{(\alpha)}\boldsymbol{\epsilon}^{(\alpha')} - \frac{1}{m}\sum_I\left[\frac{(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{BI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{IA}}{E_I - E_A - \hbar\omega - i\Gamma_I/2} + \frac{(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{BI}(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{IA}}{E_I - E_A + \hbar\omega'}\right]\right|^2. \tag{2.188}$$

In practice Γ_I can be ignored except when $E_I - E_A \simeq \hbar\omega$.

In general, the resonant amplitude is much larger than the sum of the nonresonant amplitudes. This is because the magnitude of the resonant amplitude is of the order of $\lambda\!\!\!^{-} = c/\omega$ (cf. Problem 2–8) while the magnitude of the nonresonant amplitude is of the order of r_0. Ignoring the nonresonant amplitudes, we can obtain a single-level resonance formula applicable to the scattering of light in the vicinity of a nondegenerate resonance state:

$$\frac{d\sigma}{d\Omega} = r_0^2\left(\frac{\omega'}{\omega}\right)\left(\frac{1}{m^2}\right)\frac{|(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{BR}|^2\,|(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{RA}|^2}{(E_R - E_A - \hbar\omega)^2 + \Gamma_R^2/4}. \tag{2.189}$$

It is amusing to note that this expression is identical with the probability of finding state R formed by the absorption of photon $(\mathbf{k}, \alpha)$, multiplied by the spontaneous emission probability per solid angle for $R \to B$ with the emission of photon $(\mathbf{k}', \alpha)$, and divided by the flux density. To prove this statement we merely note that from (2.184) and (2.119) we obtain

$$\frac{\text{(absorption probability)} \times \text{(emission probability per solid angle)}}{\text{flux density}} = \frac{\dfrac{e^2\hbar}{2\omega m^2 V}\left[\dfrac{|(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha)})_{RA}|^2}{(E_R - E_A - \hbar\omega^2) + \Gamma_R^2/4}\right]\left[\dfrac{2\pi}{\hbar}\dfrac{e^2\hbar}{2m^2\omega' V}|(\mathbf{p}\cdot\boldsymbol{\epsilon}^{(\alpha')})_{BR}|^2\dfrac{V\omega'^2}{(2\pi)^3\hbar c^3}\right]}{c/V}, \tag{2.190}$$

which is identical to (2.189). Thus resonance scattering can be visualized as the formation of the resonance state R due to the absorption of the incident photon followed by the spontaneous emission of the outgoing photon.

The above simple interpretation of resonance fluorescence as the formation of resonance followed by its decay is subject to the usual peculiarities of quantum mechanics. First of all, if the nonresonant amplitudes are appreciable, (2.188) tells us that they can interfere with the resonant amplitude. Interference between the nonresonant amplitudes and the resonant amplitude is frequently observed in nuclear resonance fluorescence. Further, if the resonant state has spin or, more generally, if there are several resonant states of the same or approximately the same energy, it is important to sum over the amplitudes corresponding to the various resonant states as indicated in (2.186) before squaring the amplitude. This point will be illustrated in Problem 2–8.

These considerations naturally lead to the following question: How can we decide whether a particular physical phenomenon is best described as a single quantum-mechanical process of resonance scattering or as two independent quantum-mechanical processes of absorption followed by emission? The answer to this

question depends critically on how the lifetime of the metastable state compares with the collision time, that is, the time interval during which the atom is exposed to the incident beam. To illustrate this point, let us first consider an experimental arrangement in which the temporal duration of the incident photon beam is short compared to the lifetime of the resonance. The energy resolution of such a photon beam is necessarily ill-defined, $\Delta(\hbar\omega) \gg \Gamma_R$, because of the uncertainty principle. It is then possible to select the excited resonant state by observing the exponential decay of the resonance long after the irradiation of the atom has stopped. In such a case, it is legitimate to regard the formation of the metastable state and its subsequent decay as *two independent* quantum-mechanical processes.

On the other hand, the situation is very different with scattering of a photon beam of a very well-defined energy. If the energy resolution of the incident beam is so good (or if the lifetime of the resonant state is so short) that the relation $\Delta(\hbar\omega) \ll \Gamma_R$ is satisfied, it is not possible to select the excited resonant state by a delayed lifetime measurement or any other measurement without disturbing the system. This is because the uncertainty principle demands that such a monochromatic photon beam must last for a time interval *much longer* than the lifetime of the resonance. We can still determine the lifetime of the resonance by measuring the decay width Γ_R through the variation of the scattering cross section as a function of ω, using *monochromatic* beams of *variable* energies. Under such a circumstance the interference effect between the nonresonant amplitudes and the resonant amplitude discussed earlier can become important. In such a case, it is meaningless to ask whether or not a particular photon has come from the resonance state. If we are considering an everlasting, monochromatic, incident beam, resonance fluorescence must be regarded as a single quantum-mechanical process so long as the atom is undisturbed, or so long as no attempt is made to determine the nature of the intermediate states.

In nuclear physics it is sometimes possible to obtain a sufficiently monochromatic photon beam so that we can study the Lorentzian energy dependence of the cross section in the vicinity of a resonance whose lifetime is $\lesssim 10^{-10}$ sec. For a resonance with lifetime $\gtrsim 10^{-10}$ sec, the mean life may be more readily determined by observing the exponential decay. For a nuclear level with lifetime $\sim 10^{-10}$ sec (corresponding to a decay width of about 10^{-7} eV) we can determine the lifetime both by a delayed lifetime measurement and by a study of the energy dependence of the cross section. Whenever both procedures are available, the agreement between the lifetimes determined by the two entirely different methods is quite satisfactory.

2–7. DISPERSION RELATIONS AND CAUSALITY

Real and imaginary parts of the forward scattering amplitude. In this section we shall discuss the analytic properties of the amplitude for the elastic scattering of photons by atoms in the forward direction, with no change in polarization. We defined the coherent forward scattering amplitude $f(\omega)$ by

$$(d\sigma/d\Omega)_{\theta=0,\,\epsilon^{(\alpha)}=\epsilon^{(\alpha')}} = |f(\omega)|^2. \tag{2.191}$$