FOREWORD

The lectures on which these notes are based were intended to serve as an elementary introduction to quantum optics. They were begun, for that reason, with discussions of classical experiments, and the introduction of quantum mechanical ideas was carried out fairly gradually. The most advanced knowledge of quantum electrodynamics which they require at any point is some acquaintance with the connection between the quantization of harmonic oscillators and that of fields. This is material which is covered in the first two of Professor Kroll's lectures, or in the initial chapters of a number of elementary texts on field theory.

These notes are derived from a set which was prepared from tape recordings and edited by R. Pratesi, L. Narducci, D. Forster, U. Titulaer, and P. Kelley. The author is most grateful for their generous help.

From Proceedings
"Quantum Optics and Electrodynamics"
(Les Houches 1964)
(Gordon and Breach, 1965)
editors: C. de Witt,
A. Blandin,
C. Cohen-Tannoudji

R. J. Glauber
Harvard University

Lecture I. INTRODUCTION

The field of optics, after seeming to have reached a sort of maturity, is beginning to undergo some rapid and revolutionary changes. These changes are connected with things which we have, as a matter of principle, known about for many years, but the extent to which we could put our knowledge into practice has, until just a few years ago, been extremely limited. Thus the electromagnetic character of light waves has been familiar knowledge since the last century. A vast body of theory and technique concerning the generation of electromagnetic waves has been built up during these years, but virtually all of it has dealt with radio frequency fields. Light waves of course, are of the same electromagnetic character as radio waves. But because the only ways we had of generating them in the past were extremely clumsy (in a sense we shall presently discuss at some length) there has been very little occasion until recently to apply the insights of radio-frequency theory in optics. A simple physical reason, as we shall see, lies at the bottom of this: all of the traditional types of optical sources possess a certain chaotic quality in common. They are what a radio engineer would refer to as noise generators, and all of the delicate and ingenious techniques of optics are exercises in the constructive use of noise. The invention of the optical maser has removed this barrier with almost a single stroke. It allows us to presume that we will some day be able to control fields oscillating at optical or higher frequencies with the same sort of precision and versatility that have become familiar in radio frequency technology.

Another recent change is the development of detectors which respond strongly to individual quanta of light. These have permitted us to explore the corpuscular character of optical fields. All of the traditional optical experiments have not only dealt with extremely crude sources, but have paid very little attention to the detection of individual light quanta. The detectors used were typically sensitive only to substantial numbers of photons and were quite slow in action so that we measured only intensities which had been averaged over relatively long periods of time. The new light detectors enable us to ask more subtle questions than just ones about average intensities; we can, for example, ask questions about the counting of pairs of quanta, and can make measurements of the probability that the quanta are present at an arbitrary pair of space points, at an arbitrary pair of times.

If the instrumentation in optics has made long strides in the direction of dealing with photons, it is worth mentioning that the instrumentation in the radio frequency field is leading in that direction as well. The energies of radio frequency

photons are extremely small, much smaller than the thermal fluctuation energy κT (T = noise temperature ~ room temperature for most amplifiers). There has consequently not been much need in radio frequency technology to date to pay attention to the corpuscular structure of the field. The recent invention, however, of low noise amplifiers, such as the microwave maser, has lowered the noise temperature of the detecting device to such a degree that with further progress it seems not impossible that individual photons may be detected. So, even in the microwave region, there is now a certain amount of attention being paid to the corpuscular structure of light.

It is interesting, in any case, to investigate the corpuscular nature of electromagnetic fields, because it will set the ultimate limitation to the possibility of transmitting information by means of fields. We will not discuss information theory in these lectures, but we will have some things to say which are related to noise theory. Noise theory is the classical form of the theory of fluctuations of the electromagnetic field and is quite naturally related to the theory of quantum fluctuations of the field. All of these subjects fall under a general heading which we might call photon statistics. Coherence theory too, is properly speaking, a rather small area of the same general subject. Its purpose is simply to formulate some useful ways of classifying the statistical behavior of fields.

The problem to which we shall address ourselves in these lectures is the construction of a fairly rigorous and general treatment of the problems of photon statistics. There is no need, in doing it, to make any material distinction between radio frequency and optical fields (or between these and X-ray fields for that matter). A part of the formalism, that which has to do with the definition of coherence, is suggested in fact as a way of unifying the rather different concepts of coherence, which have characterized these areas in the past.

We have already remarked that optical experiments have only rarely dealt with individual photons. Much the same observation can be made for optical theory as well. If the photon has to such a remarkable degree remained a stranger to optical theory some justification for that fact surely lies in the great success of the simple wave models in the analysis of optical experiments. Such models are usually spoken of as being classical in character since they proceed typically from some kind of analogy to classical electromagnetic theory and pay as little attention to the corpuscular character of the radiation as the experimental arrangement will permit.

In these approaches one talks typically about some kind of "optical disturbance function" which is assumed to obey the wave equation and perhaps certain boundary conditions as well. The function may represent the components of the electric vector or possibly other field quantities such as the vector potential, or the magnetic field. In many applications in fact one does not need to be very specific about what it really does or does not represent.

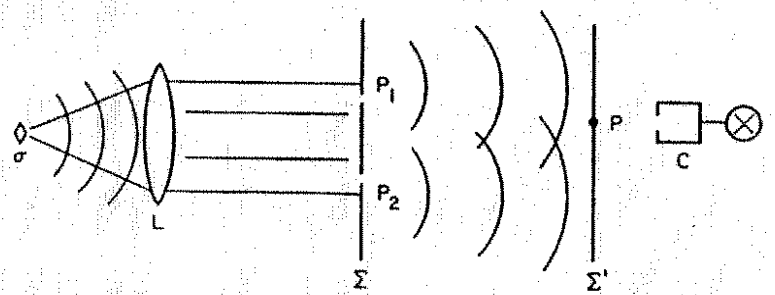

Figure 1

Let us consider the Young interferometer (Fig. 1) in order to illustrate the elementary approaches we are discussing. A plane, quasi-monochromatic wave coming from a point source σ impinges on the screen Σ with two parallel slits at the positions P_1 and P_2.

The two waves emerging from the slits give rise to an interference pattern on the screen Σ', which we can often see with the unaided eye. The simplest way of predicting the form of the interference pattern is to ignore the vector character of the electromagnetic field and introduce a scalar field φ which is presumed to describe the "optical disturbance." We then try to find a function φ which satisfies the wave equation together with a set of boundary conditions which we take to represent the effect of the screen Σ. That problem, as you remember, is in general a good deal too difficult to be solved exactly, and it is customary to make a number of simplifying approximations such as dealing very crudely with the boundary conditions, and making use of the Huyghens principle. By these familiar methods we reach a simple evaluation of the field distribution φ on the screen Σ'.

Of course, if we are to predict the form of the interference pattern, we must at some stage face the question of attaching a physical interpretation to the field φ. The most familiar approach is to regard $\varphi(\mathbf{r},t)$ as a real field and to identify it, perhaps, with one of the components of the electric field vector. The experimental fringe pattern is then predicted quite accurately, as we all know, if the light intensity on the screen is identified with φ^2, the square of our optical field. The identification possesses the justification, from the standpoint of classical theory, that the Poynting vector, which tells us the energy flux, is indeed quadratic in the field strength. In spite of this evident support the identification is not a unique one, however; it pays too little attention to the way in which the light is detected.

Let us suppose that the light intensity is measured by using a photon counter at the position of the screen. We then ask how we may predict the response of the counter as it is used to probe the pattern. Although the use of the wave equation to find the field amplitude φ did not introduce any distinctions between the classical and the quantum theoretical approaches to the diffraction problem, the use of a photon counter as a detector does introduce a distinction. The photon counter is an intrinsically quantum mechanical instrument. Its output is only predictable in terms of statistical averages even when the state of the field is specified precisely. If we are to predict this average response we must be rather more specific than we have thus far been about the field which the counter sees and we must treat the detection mechanism in a fully quantum mechanical way. What we find when we do these things is that the counter may be more accurately thought of as responding to a complex field φ^+ rather than the real field φ, and as having an output proportional, not to φ^2, but to $|\varphi^+|^2$, (The distinction is not a trivial one physically, since in a monochromatic field φ^2 oscillates rapidly in magnitude while $|\varphi^+|^2$ remains constant.) Once this answer is known it can be used as a crude rule for bypassing the explicit discussion of the detection mechanism in applications to other detection problems.

The use of such rules as a means of avoiding the explicit use of quantum mechanics has several times been called the "semi-classical approach". While approaches of this type clearly need a rule of some sort to bridge the gap between their descriptions of the wave and particle behaviors of photons they may remain perfectly correct approaches in a quantum mechanical sense as long as the rule has been chosen correctly. The fact that a mistaken form of this rule has been used repeatedly in "semi-classical" discussions is a good indication that the fully quantum mechanical discussion is not entirely beside the point.

One of the properties of the "semi-classical" approaches that makes them elementary is that they deal with ordinary numbers and functions. They make no use of the apparatus of non-commuting operators which, it may appear, ought to be part of any formal quantum mechanical description of the field. Later in these

lectures we shall show that for a certain class of fields there need be no error in a statistical description of the field which is based upon such ordinary functions as we find by solving the wave equation. It is possible to describe these fields fully by means which are rather similar to those used in the classical theory of noise. Where such a description is available it means that there need be nothing incorrect about the so-called "classical" or "semi-classical" approaches except their names, which then become totally misleading. It has recently been claimed that the class of states of the field for which the simple statistical description we have mentioned is available includes all states of the field, and that consequently the quantum theory and the "classical" theory will always yield equivalent results. We shall have to return to this point later in the lectures when we are better equipped to discuss it, but for the present we may remark that this claim seems to be based more upon wishful thinking than upon accurate mathematics. The quantum theory still offers the only complete and logically consistent basis for discussing field phenomena.

The general subject we shall be discussing, to give it its most imposing name, is quantum electrodynamics. It is an extremely well developed subject. Although it has long been clear that classical electrodynamics is the limit of quantum electrodynamics for $\hbar \to 0$, there have never been any very powerful methods available for discussing electrodynamical problems near the classical limit.

All of quantum electrodynamics has historically been developed in terms of the stationary states $|\, n \,\rangle$ of the field hamiltonian $\mathcal{H}$. These correspond to the presence of an integer number n of quanta, i. e. they obey the equation

$$\mathcal{H} \,|\, n \,\rangle = (n + \tfrac{1}{2})\, \hbar\, \omega \,|\, n \,\rangle \,. \tag{1.1}$$

The n-quantum states form a complete set which has usually been regarded as the "natural" basis for the development of all states of the field. To the extent that virtually all electrodynamical calculations have been done by means of expansions in powers of the field strengths, the numbers of photons which have been dealt with in the calculations have usually been very small integers. The classical limit of quantum electrodynamics, on the other hand, is one in which the quantum numbers are typically quite large. Not only are they large but they are typically quite uncertain. If, for example, a harmonic oscillator is vibrating in a state with a relatively well defined phase, it is necessary that it not only be in a state with a large quantum number, but that the quantum number of the state also be quite uncertain, $(\Delta n\, \Delta\varphi \gtrsim 1)$. When we must deal with quantum states of the electromagnetic field for which the phase of the field is well defined, they can likewise only be states in which the occupation number n is intrinsically rather indefinite. In such cases the description of expectation values in terms of the n- quantum states becomes rather awkward and untransparent.

One of the mathematical tools we shall use in these lectures is a set of quantum states rather better suited to the description of amplitude and phase variables than the n-quantum states. The use of these states makes the relationship of the classical and quantum mechanical forms of electrodynamics considerably clearer than it has been before.

CLASSICAL THEORY

It may help to underscore the close connection between the quantum theory we shall develop and the classical theory if we begin by discussing the classical theory alone for a while. We shall describe the classical field in terms of the familiar field vectors, the electric field $\mathbf{E}(\mathbf{r}, t)$ and the magnetic field $\mathbf{B}(\mathbf{r}, t)$. We will take these to obey the source-free Maxwell equations

$$\nabla \cdot \mathbf{E} = 0, \qquad \nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}$$
$$\nabla \cdot \mathbf{B} = 0, \qquad \nabla \times \mathbf{B} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} \tag{1.2}$$

by assuming that whatever source has radiated the fields has ceased to radiate further.

Since our detectors are usually sensitive to electric rather than magnetic fields, we shall confine ourselves to a discussion of the field $\mathbf{E}(\mathbf{r}, t)$. One of the first things which is done in many classical calculations is to use a Fourier series or integral to expand the time dependence of the field and in that way to separate the field into two complex terms:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}^{(+)}(\mathbf{r}, t) + \mathbf{E}^{(-)}(\mathbf{r}, t) \tag{1.3}$$

The first of these terms, which we shall call the positive frequency part, $\mathbf{E}^{(+)}$, contains all the amplitudes which vary as $e^{-i\omega t}$ for $\omega > 0$. The other, the negative frequency part, contains all amplitudes which vary as $e^{i\omega t}$. These terms are complex conjugates of one another

$$\mathbf{E}^{(-)} = \mathbf{E}^{(+)\dagger} \tag{1.4}$$

and contain equivalent physical information. Either one or the other is frequently used in classical calculations and called either the complex field strength or the complex signal. The use of these complex fields in classical contexts is usually regarded as a mathematical convenience rather than a physical necessity since classical measuring devices tend to respond only to the real field, $\mathbf{E} = 2\,\mathrm{Re}\,\mathbf{E}^{(+)}$.

Quantum mechanical detectors, as we have noted, behave rather differently from classical ones, and for the discussion of these the separation of the field into its positive and negative frequency parts takes on a much deeper significance than it does for classical detectors. As we shall later see, an ideal photon counter (one which has zero size and is equally sensitive to all frequencies) measures the product $\mathbf{E}^{(-)}(\mathbf{r}, t)\, \mathbf{E}^{(+)}(\mathbf{r}, t) = |\mathbf{E}^{(+)}(\mathbf{r}, t)|^2$. That, at least, is what the detector would measure if we were capable of preparing fields with precisely fixed field strengths. But of course we are never capable of controlling the motions of the charges in our sources with very great precision. In practice all fields are radiated by sources whose behavior is subject to considerable statistical uncertainty. The fields are then correspondingly uncertain and what we require is a way of describing this uncertainty in mathematical terms.

It is more convenient, in describing the randomness of the fields, to deal with a discrete set of variables than to deal with the whole continuum at once. We shall therefore only attempt to describe the field lying inside a certain volume of space within which we can expand it in terms of a discrete set of orthogonal mode functions. We shall take the set of vector mode functions $\{\mathbf{u}_k(\mathbf{r})\}$ to obey the wave equations

$$\left(\nabla^2 + \frac{\omega_k^2}{c^2}\right) \mathbf{u}_k(\mathbf{r}) = 0\,, \tag{1.5}$$

which define a set of frequencies $\{\omega_k\}$ when they are satisfied together with the constraint

$$\nabla \cdot \mathbf{u}_k(\mathbf{r}) = 0 \tag{1.6}$$

and a suitable set of boundary conditions. These functions may be assumed to form an orthonormal set

$$\int \mathbf{u}_k^*(\mathbf{r}) \cdot \mathbf{u}_{k'}(\mathbf{r})\, d\mathbf{r} = \delta_{kk'} \tag{1.7}$$

which is complete within the volume being studied. They may then be used to express the electric field vector in the form

$$E(r,t) = \Sigma_k C_k u_k(r) e^{-i\omega_k t} + \Sigma_k C_k^* u_k^*(r) e^{i\omega_k t}, \qquad (1.8)$$

The two sums on the right are then evidently $E^{(+)}$ and $E^{(-)}$, respectively.

When the expansion in orthogonal modes is used the field is evidently specified completely by the set of complex Fourier amplitudes $\{C_k\}$. To describe random fields we must regard these numbers as random variables in general. Usually the most we can state about these coefficients can be expressed through a probability distribution $p(\{C_k\}) = p(C_1, C_2, C_3, \ldots)$. Then, if we measure some function of E or of $E^{(+)}$, the most we can hope to predict is its mean value, i.e., if we measure $F(E^{(+)})$ we can only hope to find the average

$$\langle F(E^{(+)})\rangle = \int p(\{C_k\})\, F[E^{(+)}(\{C_k\})] \Pi_k d^2 C_k , \qquad (1.9)$$

where the differential element of area is given by $d^2C_k = d(\mathrm{Re}C_k)\, d(\mathrm{Im}C_k)$.

It is important to remember that this average is an ensemble average. To measure it we must in principle repeat the experiment many times by using the same procedure for preparing the field over and over again. That may not be a very convenient procedure to carry out experimentally but it is the only one which represents the precise meaning of our calculation. The fields we are discussing may vary with time in arbitrary ways. As an example we might take the field generated by a radio transmitter sending some arbitrarily chosen message. There is therefore no possibility in general of replacing the ensemble averages by time averages. The theory of non-stationary statistical phenomena can only be developed in terms of ensemble averages.

The solution of problems in statistical thermodynamics has accustomed us to thinking of statistical fluctuations about the ensemble average as being very small. We are thus usually willing to forget about the need in principle to make an ensemble of thermodynamic measurements and are content to compare just a single measurement with the predicted ensemble average. While the justification of such shortcuts may be excellent in thermodynamic contexts, it is not always so good in statistical optics. Thus when we speak later of the interference patterns produced by superposing light from independent sources we shall find that individual measurements yield results wildly unlike their ensemble averages. The distinction between particular measurements and their averages may thus be quite essential.

Lecture II. INTERFERENCE EXPERIMENTS

One of the classic experiments which exhibits the coherence properties of light is the Young experiment (Fig. 2).

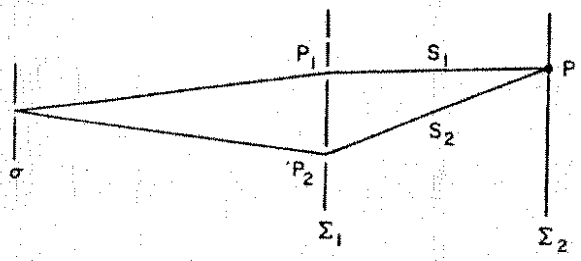

Figure 2

The field present at P at time t may be approximated by a certain linear superposition of the fields present at the two pinholes at earlier times:

$$E^{(+)}(r,t) = \lambda_1 E^{(+)}(r_1, t_1) + \lambda_2 E^{(+)}(r_2, t_2) \qquad (2.1)$$

where the times are given by $t_{1,2} = t - S_{1,2}/c$. The coefficients λ_1, λ_2 depend on the geometry of the arrangement, but are taken to be independent of the properties of the field.

We shall assume, to begin the discussion, that a photodetector placed at P measures the squared absolute value of some component of the complex field strength. (At a later point we shall discuss the validity of the assumption in some detail.) If we write the measured field component as $E^{(+)}(r,t)$, we then have

$$\begin{aligned} |E^{(+)}(r,t)|^2 = E^{(-)}(r,t) E^{(+)}(r,t) = |\lambda_1|^2 E^{(-)}(r_1,t_1) E^{(+)}(r_1,t_1) \\ + |\lambda_2|^2 E^{(-)}(r_2,t_2) E^{(+)}(r_2,t_2) \\ + 2\,\mathrm{Re}\{\lambda_1^* \lambda_2 E^{(-)}(r_1,t_1) E^{(+)}(r_2,t_2)\}. \end{aligned} \qquad (2.2)$$

Now since our preparation of the source rarely fixes the Fourier coefficients C_k very precisely we must in principle perform the experiment repeatedly and then average in order to find a non-random result. The only thing we can really predict is the ensemble average of $|E^{(+)}(r,t)|^2$ taken over the set of random coefficients $\{C_k\}$,

$$\begin{aligned} \langle |E^{(+)}(r,t)|^2\rangle = |\lambda_1|^2 \langle |E^{(+)}(r_1,t_1)|^2\rangle + |\lambda_2|^2 \langle |E^{(+)}(r_2,t_2)|^2\rangle \\ + 2\mathrm{Re}\, \lambda_1^* \lambda_2 \langle E^{(-)}(r_1,t_1) E^{(+)}(r_2,t_2)\rangle . \end{aligned} \qquad (2.3)$$

If we introduce the first order correlation function

$$G^{(1)}(rt, r't') = \langle E^{(-)}(rt) E^{(+)}(r't')\rangle , \qquad (2.4)$$

we can rewrite Eq. (2.3) in the following way

$$\begin{aligned} \langle |E^{(+)}(r,t)|^2\rangle = |\lambda_1|^2 G^{(1)}(r_1t_1, r_1t_1) + |\lambda_2|^2 G^{(1)}(r_2t_2, r_2t_2) \\ + 2\mathrm{Re}\{\lambda_1^* \lambda_2 G^{(1)}(r_1t_1, r_2t_2)\}, \end{aligned} \qquad (2.5)$$

We have omitted consideration of vector and tensor indices of the fields and correlation functions, respectively, since the vector properties of the field are not too important in this experiment. We would have to take careful account of them if somehow a rotation of the plane of polarisation were induced behind one pinhole, or if the polarisation were in any way made to play a more active role.

A particular case which occurs almost universally in classic optics is that in which the incident field is stationary. The term "stationary" does not mean that nothing is happening. On the contrary, the field is ordinarily oscillating quite rapidly. It means that our knowledge about the field does not change with time. More formally, we associate stationarity with invariance of the statistical description of the beam under displacements of the time variable. The correlation function $G^{(1)}$ for such fields can therefore only depend on the difference $t - t'$

$$G^{(1)}(t, t') = G^{(1)}(t - t') \quad \text{(stationary field)}. \qquad (2.6)$$

(Note that by discussing only a single type of correlation function we are stating a necessary condition for stationarity, but not a sufficient one. All average properties of a stationary field must be unchanged by time displacements.) When random classical fields are represented by means of stationary stochastic processes the models used usually have the ergodic property. That property means that the function $G^{(1)}(t - t')$ which is defined as an ensemble average, has the same value as the time averaged correlation function $\Gamma^{(1)}(t - t')$,

$$G^{(1)}(r_1, r_2, \tau) = \Gamma^{(1)}(r_1, r_2, \tau) = \lim_{T\to\infty} \frac{1}{T}\int_0^T E^{(-)}(r_1, t_1+\tau)\, E^{(+)}(r_2, t_1)\, dt_1 . \tag{2.7}$$

The properties of the time-averaged correlation functions $\Gamma^{(1)}$ for classical fields have been discussed in detail in Chapter X of the text of Born and Wolf.

It may be of some help in the lectures that follow to have some more concrete applications of interference experiments in mind. Let us take a brief look at one of the fundamental techniques of interferometry by considering a case in which the field incident upon a detector is a superposition of two plane waves. We assume that the propagation vectors of the two plane waves are only very slightly different. This might be the case for example for monochromatically filtered light from the two members of a double star. If we assume that the frequencies of both waves are equal we may write

$$E^{(+)}(r, t) = A\, e^{i(k\cdot r-\omega t)} + B\, e^{i(k'\cdot r-\omega t)} , \tag{2.8}$$

The question we now ask is: what kind of measurement can be performed to determine that we are receiving radiation from two sources and not just one?

Before answering the question let us specify the statistical character of the coefficients A and B. They are, of course, particular examples of the coefficients C_k previously introduced. We will assume A and B to be distributed independently of one another. This means that the probability function $p(A, B)$ factorizes,

$$p(A, B) = p_1(A)\, p_2(B) , \tag{2.9}$$

We will assume further more as properties of the distributions p_1 and p_2, that the phases of the complex amplitudes A and B are individually random. We then have $\langle A\rangle = \langle B\rangle = 0$. More generally the mean values of various powers of the amplitudes and their complex conjugates such as $\langle A\, B^*\rangle$, $\langle |A|^2\, A^*\, B\rangle$, etc. will vanish. Averages in which the amplitudes are paired with their complex conjugates however, take on positive values,

$$\langle |A|^{2n}\rangle \neq 0, \qquad \langle |B|^{2n}\rangle \neq 0, \qquad n = 1, 2 \ldots \tag{2.10}$$

A famous device invented to answer the question we have asked is the Michelson stellar interferometer (Fig. 3).

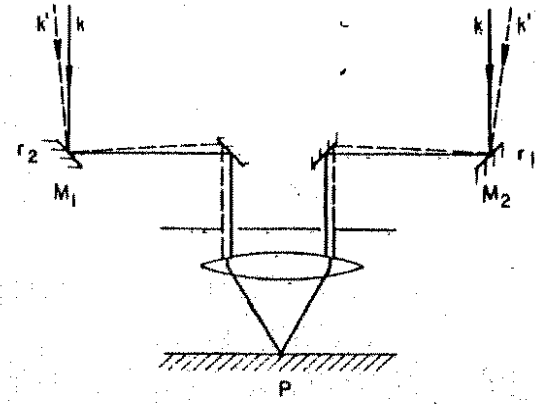

Figure 3

The field at the point P and time t is, in effect, the sum of the two fields impinging on the mirrors M_1, M_2 at the same instant t' (if the optical paths $\overline{M_1 P}$ and $\overline{M_2 P}$ are equal). Each of these two fields is of the form (2.8) evaluated at the points r_1 and r_2 respectively. The average intensity at P will therefore be

$$\langle E^{(-)}(r, t)\, E^{(+)}(r, t)\rangle = 2\,\mathrm{Re}\{\langle |A|^2 + |B|^2\rangle + \langle |A|^2\rangle e^{-ik\cdot(r_1-r_2)} + \langle |B|^2\rangle e^{-ik'\cdot(r_1-r_2)}\}, \tag{2.11}$$

where we have used $\langle AB^*\rangle = \langle A\rangle\langle B^*\rangle = 0$ in reaching this expression. If we introduce the correlation function (2.4),

$$G^{(1)}(r_1 t', r_2 t') = \langle E^{(-)}(r_1 t')\, E^{(+)}(r_2 t')\rangle = \langle |A|^2\rangle e^{-ik\cdot(r_1-r_2)} + \langle |B|^2\rangle e^{-ik'\cdot(r_1-r_2)}, \tag{2.12}$$

then intensity may be written as

$$\langle E^{(-)}(r, t)\, E^{(+)}(r, t)\rangle = 2\mathrm{Re}\{\langle |A|^2 + |B|^2\rangle + G^{(1)}(r_1 t', r_2 t')\}. \tag{2.13}$$

The correlation function which describes the interference effect is time independent, because of the stationary character of the field we are treating.

We see from Eq. (2.12) that the correlation function contains two spatially oscillating terms. The way in which these terms reinforce or cancel one another will depend on the displacement $r_1 - r_2$. If $\langle |A|^2\rangle = \langle |B|^2\rangle$, Eq. (2.11) yields

$$\langle E^{(-)}(r, t)\, E^{(+)}(r, t)\rangle = 4\langle |A|^2\rangle\{1 + \cos[\tfrac{1}{2}(k+k')\cdot(r_1-r_2)]\times \cos[\tfrac{1}{2}(k-k')\cdot(r_1-r_2)]\} \tag{2.14}$$

The interference intensity which we see at the point r will be part of a pattern of parallel fringes which we see at the focus of the telescope. Although we have not attempted to describe the fringe pattern in detail, the expression (2.14) for the intensity does indicate one of the characteristic properties of the pattern, that it will vanish altogether when the displacement $r_1 - r_2$ is adjusted so that

$$\cos[\tfrac{1}{2}(k-k')\cdot(r_1-r_2)]$$

passes through the value zero. By observing the fringes we know that we are dealing with two sources rather than one, and by finding the values of $r_1 - r_2$ at which the fringes disappear we determine their angular separation. The Michelson interferometer has indeed been used to measure the angular separations of double stars, and for measuring angular diameters of stars as well. Only a few stellar diameters have been measured in this way, however, because of the difficulties inherent in working with a large interferometer. An unusually great mechanical stability is clearly required of the apparatus. Furthermore random variations of the index of refraction along the optical path can wash out the pattern.

Instruments quite similar to the Michelson stellar interferometer have been used in radio-astronomy to determine the angular size of celestial radio sources. They consist of two separated antennas supplying signals to a common detector system. In the case of these instruments, as well, it is technically difficult to increase the separation of the antennas without introducing random phase differences in the path between the antennas and the detector. To overcome this difficulty Hanbury Brown and Twiss have devised another form of radio interferometer (Fig. 4).

The signals at the antennas are detected individually and then the detector outputs, which are of much lower frequency, are transmitted to a central correlating device where they are multiplied together and the product is averaged. The angular size of the source is obtained from measurements of the way in which the correlation of the intensity fluctuations of the signals varies with the separation of the antennas. An equivalent arrangement may be used with visible light.

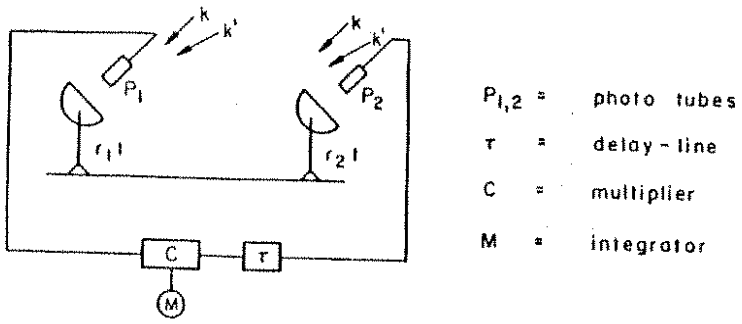

Figure 4

The essence of the trick used by Hanbury Brown, and Twiss was to detect the signals first and by taking away the high frequency components of the incoming radiation, to transmit to the central observation point just a measure of the fluctuations of the intensities arriving at the receivers. Since the detector signals are of relatively low frequency they are easy to transmit faithfully over distances large compared to the limiting dimensions of Michelson interferometers. This experiment is quite different in nature from the interferometer experiment we described earlier because it deals with the average of the product of two random intensities rather than with a single intensity.

It is easy to see that in the average of the product of the two signals there is an interference term, which permits us to resolve the two incoming waves. First we note that a square-law detector placed at P_1 gives a response proportional to

$$|E^{(+)}(r_1,t)|^2 = |A|^2 + |B|^2 + AB^* e^{i(k-k')\cdot r_1} + A^*B\, e^{-i(k-k')\cdot r_1} , \qquad (2.16)$$

This output no longer contains the rapid oscillations of the incoming wave. An average of this detected signal, however, would have no interference term (since $\langle AB^* \rangle = 0$). What Hanbury Brown, and Twiss did is multiply together the two detected signals and then, and only then, to measure the statistical average. The average of the product of two intensities of the form of Eq. (2.16) is

$$\langle |E^{(+)}(r_1t)|^2 \, |E^{(+)}(r_2t)|^2 \rangle = \langle (|A|^2 + |B|^2)^2 \rangle + 2\langle |A|^2 |B|^2 \rangle \cos\left[(k-k')\cdot(r_1 - r_2)\right], \qquad (2.17)$$

where we have used the fact that $\langle |A|^2 A^* B \rangle = 0$, etc. The cosine term clearly represents an interference effect. We can use it to resolve the two sources by observing its behavior as $r_1 - r_2$ is varied. It is important to note that the interference effect has been found by considering the average of a quantity quartic in the field amplitudes. In the case of Michelson's interferometer we deal only with expressions quadratic in the field amplitudes.

Although we have discussed the interferometer experiments in terms of ensemble averages, it is clear that they are not ordinarily performed in this way, but rather as time averages. The calculation of time averages, however, is typically at least a little more difficult than the calculation of ensemble averages (and often it is incomparably more difficult). To consider the interferometer measurements as time averages we should have to note that the two plane waves are not, in general, perfectly monochromatic. It follows then that the coefficients A and B, which we were content earlier to evaluate only at a particular instant of time, actually vary with time. To proceed further we should have to adopt models to represent A(t) and B(t) as stochastic functions of time. As we shall see presently, there are extremely persuasive reasons, when we are dealing with natural light sources, to take these models to be Gaussian stochastic processes. Then, since such processes have the ergodic property, we are justified in identifying time averages with ensemble averages.

REFERENCE

M. Born and E. Wolf, Principles of Optics (Pergamon Press, Inc., London, 1959), Chap. X.

Lecture III INTRODUCTION OF QUANTUM THEORY

When we describe the electromagnetic field in quantum mechanical terms we must think of the field vectors E and B as operators which satisfy the Maxwell equations. The states, $|\,\rangle$, on which these operators act and their adjoints, $\langle\,|$, contain the information which specifies the field. When measurements are made of the physical quantity which correspond to an operator $\mathcal{O}$, we can not expect in general to find the same results repeatedly. What we find instead is that the measured values fluctuate about the average value given by the product $\langle\,|\mathcal{O}|\,\rangle$. The fluctuation is only absent if the state, $|\,\rangle$, happens to be an eigenstate of $\mathcal{O}$, i.e., if we have

$$\mathcal{O}|\,\rangle = \mathcal{O}'|\,\rangle, \qquad (3.1)$$

where $\mathcal{O}'$ is an ordinary number rather than an operator. In that case it is convenient to use Dirac's convention and let the eigenvalue $\mathcal{O}'$ be a label for the state by writing the latter as $|\mathcal{O}'\rangle$.

As in classical electromagnetic theory, it is convenient to separate the field operator, $\mathbf{E}(\mathbf{r},t)$, which is naturally Hermitian, into the sum of its positive frequency and negative frequency parts:

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}^{(+)}(\mathbf{r},t) + \mathbf{E}^{(-)}(\mathbf{r},t) \qquad (3.2)$$

These parts, as we have already noted classically, represent complex rather than real fields. The operators $\mathbf{E}^{(+)}$ are therefore not Hermitian, but they are Hermitian adjoints of one another

$$\mathbf{E}^{(-)}(\mathbf{r},t) = \{\mathbf{E}^{(+)}(\mathbf{r},t)\}\dagger \, . \qquad (3.3)$$

While the fields $\mathbf{E}^{(+)}$ and $\mathbf{E}^{(-)}$ play essentially indistinguishable roles in classical theory, they tend to play quite dissimilar roles in the quantum theory. The

operator $E^{(+)}$ describes the annihilation of a photon while $E^{(-)}$ describes the creation of one. This identification of the operators is virtually the only fact we shall have to borrow from more formal developements of quantum field theory.

We must think fundamentally of all electric field measurements as being made on the Hermitian operator $E(r, t)$ given by Eq. (3.2). In the classical limit it is usually true that the complex fields $E^{(+)}$ and $E^{(-)}$ make contributions of equal magnitude to our measurements. From a quantum mechanical standpoint that is because quantum energies are so small in the classical limit ($\hbar\omega \to 0$), that test charges emit quanta as readily as they absorb them. In the quantum domain, on the other hand, we must expect that the fields $E^{(+)}$ and $E^{(-)}$ will make contributions of altogether different magnitudes to the quantities we measure, such as transition amplitudes.

If we are using atomic systems in their ground states as probes of the electric field for example, then the atoms have no energy to emit photons and can only absorb them. In this case, which corresponds in principle to that of a typical photodetector, only the annihilation operator $E^{(+)}$ figures significantly in determining the transition amplitudes. More exactly, if we do a calculation of the transition amplitude using first order perturbation theory, we easily find that the creation operator $E^{(-)}$ contributes only an extremely small amplitude which varies so rapidly with time that it leads to no observable effect at all. The creation operator can only contribute materially if the detector contains excited atoms. (Thermal energies are a great deal too small to furnish atoms excited to optical energies, but at microwave frequencies it may be necessary to take thermally excited atoms into account.)

In the third and higher orders of perturbation theory, the creation operator can indeed play a tiny role in an absorption experiment. The effect in question is a radiative correction to the first order absorption probability which all estimates indicate will be quite small. We see, therefore, that it is fairly accurate to say that a typical photodetector detects the field $E^{(+)}$ rather than the field E. Although this statement is clearly an approximate one rather than a rigorous one it is none the less important since it furnishes us a reason for formulating the theory in terms of a set of non-Hermitian operators. The formulation, as we shall see, allows in turn a great deal of insight into the way the theory passes to the classical limit.

To gain some further insights into the kinds of quantities measured in photon counting experiments, let us examine the role played by the field operator in the calculation of the appropriate transition probabilities. In the next lecture we shall indicate how these transition probabilities are calculated in some detail by taking due account of the atomic nature of the detector. Let us for the moment, however, ignore the detailed dynamics of the detector and assume simply that it is an ideally selective device, one which is sensitive to the field $E^{(+)}(rt)$ at a single point of space r at each instant of time t. We may take the transition probability of the detector for absorbing a photon from the field at position r and time t to be proportional to

$$w_{i \to f} = |\langle f| E^{(+)}(rt) |i\rangle|^2 , \tag{3.4}$$

where $|i\rangle$ is the initial state of the field before the detection process, and $|f\rangle$ is a final state in which the field could be found after the process. In fact we never measure the final state of the field. The only thing we do measure is the total counting rate. To calculate the total rate we have to sum Eq. (3.4) over all the final states of the field that can be reached from $|i\rangle$ by an absorption process. We can, however, extend the sum over a complete set of final states since the the states which cannot be reached (e.g., states $|f\rangle$ which differ from $|i\rangle$ by two or more photons) simply will not contribute to the result since they are orthogonal to the state $E^{(+)}(rt)|i\rangle$.

When the final state summation is carried out the counting rate becomes, in effect,

$$w = \sum_f |\langle f| E^{(+)}(r, t) |i\rangle|^2 = \langle i| E^{(-)}(r, t) E^{(+)}(r, t) |i\rangle , \tag{3.5}$$

where the completeness relation $\Sigma |f\rangle\langle f| = 1$ has been used. The counting rate w is proportional to the probability per unit time that an ideal photocounter, placed at r, absorbs a photon from the field at time t. It is, according to Eq. (3.5), given by the expectation value of the positive definite Hermitian operator $E^{(-)}(r,t) E^{(+)}(r, t)$, taken in the state $|i\rangle$ which the field was in prior to the measurement. Eq. (3.5) shows explicitly that the photocounter is not sensitive to the square of the real field (as has been assumed in many "semi-classical" calculations), but rather to an operator which corresponds to the squared absolute magnitude the complex field-strength.

We have thus far supposed that we know the state $|i\rangle$ of the field. That does not mean, of course, that we can predict the result of a single measurement made with our counter. If we repeat the measurement another result will quite likely turn out, and Eq. (3.5) gives us only the mean value of many repeated measurements. So quantum mechanics forces us to talk about ensemble averages even if we know the state of the field precisely.

In practice, of course, we almost never know the state $|i\rangle$ very precisely. Radiation sources are usually complicated systems with many degrees of freedom, so the states $|i\rangle$ depend, as a rule, on many uncontrollable parameters. Since we have no possibility of knowing the exact state of a field, we must resort to a statistical description. This description summarizes our knowledge of the field, by averaging over the unknown parameters. The predictions that we make by using this description must therefore, in principle, be compared experimentally with ensemble averages. With this understanding we may write the counting rate as an ensemble average of Eq. (3.5) over all random variables involved in the state $|i\rangle$,

$$w = \left\{ \langle i| E^{(-)}(r, t) E^{(+)}(r, t) |i\rangle \right\}_{\text{av. over } i} . \tag{3.6}$$

If we introduce the density operator $\rho = \left\{ |i\rangle\langle i| \right\}_{\text{av. over } i}$, we may write this average as

$$w = \mathrm{Tr} \left\{ \rho E^{(-)}(r, t) E^{(+)}(r, t) \right\} , \tag{3.7}$$

where Tr stands for the trace of the operator which follows. The density operator is the average of the projection operators on the initial field states. It is obviously Hermitian, $\rho\dagger = \rho$. Furthermore, it also has the property of positive definiteness, $\langle j|\rho|j\rangle \geq 0$ for any state $|j\rangle$. It is worth emphasizing that a two-fold averaging process is implied by Eq. (3.7). That we must average the measurements made upon a pure state is an intrinsic requirement of quantum mechanics which has no classical analogue. The ensemble average over initial states, on the other hand, is analogous to the averaging over the set of random coefficients $\{C_k\}$ which we described in the classical theory.

Equation (3.7) gives the counting rate of a single ideal photodetector in terms of the quantum mechanical correlation function

$$G^{(1)}(x, x') = \mathrm{Tr} \left\{ \rho E^{(-)}(x) E^{(+)}(x') \right\}, \qquad x \equiv \{r, t\} , \tag{3.8}$$

which is analogous to the correlation function introduced to describe classical interference experiments. To describe more sophisticated experiments, e.g., the coincidence experiment of Hanbury Brown and Twiss, it is useful to define a more general set of correlation functions

$$G^{(n)}(x_1 \cdots x_n, x_{n+1} \cdots x_{2n}) = \mathrm{Tr} \{ \rho E^{(-)}(x_1) \cdots E^{(-)}(x_n) E^{(+)}(x_{n+1}) \cdots E^{(+)}(x_{2n}) \} . \tag{3.9}$$

The function $G^{(n)}$ will be referred to as the n-th order correlation function. The analytical properties of this set of functions and their relation to experimental measurements will be discussed later.

We could, of course, have chosen to define a somewhat larger class of correlation functions than the $G^{(n)}$ by dealing with averages such as $\mathrm{Tr}\{\rho E^{(-)}E^{(+)}E^{(+)}E^{(+)}\}$, which contain unequal numbers of creation and annihilation operators. If we have chosen not to set down any special notation for such averages it is because they are not of the types which are measured in typical photon counting experiments. Such averages may, in principle, be measured in other kinds of experiments but they will always vanish in stationary states of the field and, much more generally, whenever the absolute phases of the fields are random. Random absolute phases are, of course, rather characteristic of optical and other extremely high frequency fields.

Lecture IV THE ONE-ATOM PHOTON DETECTOR

Let us now consider the photodetection process in somewhat more detail. We shall imagine, for the present, that our photon counter is a rather idealized type of device which has as its sensitive element a single atom which is free to undergo photoabsorption transitions such as the photoelectric effect. We assume that the atom is shielded from the radiation field we are investigating by a shutter of some sort which opens at time t_0 and closes again at time t. Our problem will be to calculate the probability that a photoabsorption process takes place during this interval and that it is recorded by our apparatus.

The detector will be assumed to be far enough from the radiation source so that the field behaves as a free field. The hamiltonian of the system (field + detector) can then be written as

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1 \quad ; \quad \mathcal{H}_0 = \mathcal{H}_{0,at} + \mathcal{H}_{0,F} ,$$

where $\mathcal{H}_0$ is the sum of Hamiltonians of the free field and the atom. The interaction term $\mathcal{H}_1$ is time independent in the Schrödinger picture. In the interaction representation, however, it becomes time dependent. If we make use of the electric dipole approximation, which is quite accurate at optical frequencies, we can write the time dependent interaction Hamiltonian as

$$\mathcal{H}_1 = e^{\frac{i}{\hbar}\mathcal{H}_0 t}\, \mathcal{H}_1\, e^{-\frac{i}{\hbar}\mathcal{H}_0 t} = -e \sum_\gamma q_\gamma(t) \cdot \mathbf{E}(\mathbf{r}, t) . \qquad (4.1)$$

In this expression r represents the position of the atomic nucleus and q_γ the position operator of the γ-th electron relative to the nucleus. The time dependence of the field $\mathbf{E}(\mathbf{r}, t)$ which occurs in Eq. (4.1) is that of the free field uninfluenced by the presence of the atom.

The Schrödinger equation of the combined system of field and atom in the interaction representation is

$$i\hbar \frac{\partial}{\partial t} |t> = \mathcal{H}_1(t) |t> . \qquad (4.2)$$

Its solution can be written in the general form

$$|t> = U(t, t_0) |t_0> ,$$

where $U(t, t_0)$ is the unitary time development operator which describes the way in which the initial state changes under the influence of the perturbation. In the first order of perturbation theory the solution has the well-known form

$$|t> = \left\{1 + \frac{1}{i\hbar} \int_{t_0}^{t} \mathcal{H}_1(t')dt'\right\} |t_0> . \qquad (4.3)$$

Let us suppose now that the system is initially in the state $|gi> = |g>|i>$, where $|i>$ is some known state of the field, and $|g>$ is the ground state of the atom. We ask now for the probability that the system at time t is in a specified state $|af> = |a>|f>$, where $|a>$ is an excited state of the atom and $|f>$ is the final state of the field. This probability is given by the squared absolute value of the matrix element

$$<af|U(t, t_0)|gi> = \frac{1}{i\hbar} \int_{t_0}^{t} <af|\mathcal{H}_1(t')|gi> dt' . \qquad (4.4)$$

(The zeroth order term in $U(t, t_0)$ of Eq. (4.3) does not contribute because of the orthogonality of the electron states $|a>$ and $|g>$.) By substituting the interaction operator from Eq. (4.1) we can separate the matrix element into two parts, a matrix element for the atom and one for the field:

$$<af|U(t, t_0)|gi> = \frac{ie}{\hbar} \sum_\gamma \int_{t_0}^{t} <a|q_\gamma(t')|g> \cdot <f|\mathbf{E}(\mathbf{r}, t')|i> dt' . \qquad (4.5)$$

To evaluate the atomic matrix element we recall that

$$q_\gamma(t') = e^{\frac{i}{\hbar}\mathcal{H}_0 t'} q_\gamma(0)\, e^{-\frac{i}{\hbar}\mathcal{H}_0 t'} = e^{\frac{i}{\hbar}\mathcal{H}_{0,at} t'} q_\gamma(0)\, e^{-\frac{i}{\hbar}\mathcal{H}_{0,at} t'} .$$

The latter relation holds because the field hamiltonian $\mathcal{H}_{0,F}$ commutes with the atomic Hamiltonian $\mathcal{H}_{0,at}$ and with the electron coordinate $q_\gamma(0)$ as well. We may write the matrix element as

$$<a|\Sigma_\gamma q_\gamma(t')|g> = M_{ag} e^{i\omega_{ag} t'}$$

with

$$M_{ag} = <a|\Sigma_\gamma q_\gamma(0)|g> \quad \text{and} \quad \hbar\omega_{ag} = E_a - E_g .$$

The matrix element M_{ag} occurs simply as a time independent coefficient in the transition amplitude

$$<af|U(t, t_0)|gi> = \frac{ie}{\hbar} \int_{t_0}^{t} e^{i\omega_{ag} t'} M_{ag} \cdot <f|\mathbf{E}(\mathbf{r}, t')|i> dt' . \qquad (4.6)$$

We can now replace $\mathbf{E}(\mathbf{r}, t')$ in this expression by the sum of the two operators $\mathbf{E}^{(+)}(\mathbf{r}, t)$ and $\mathbf{E}^{(-)}(\mathbf{r}, t)$. The emission operator $\mathbf{E}^{(-)}(\mathbf{r}, t)$ contains only negative frequencies, i.e., exponential time dependences of the form $e^{i\omega t}$ for $\omega > 0$. The time integrals of these terms clearly oscillate rapidly with increasing t. They are furthermore quite small in amplitude compared with the terms contributed by the annihilation operator $\mathbf{E}^{(+)}(\mathbf{r}, t)$. What we are describing, in fact, is the way in which the transitions are restricted by the conservation of energy. In order to find that the atomic transitions conserve the energy of the field quanta with an accuracy $\Delta E = \hbar\Delta\omega$, we must leave our shutter open for a length of time $t - t_0 \gg 1/\Delta\omega$. In practice we always have $\Delta\omega \ll \omega_{ag}$, i.e., the shutter is open for a great many periods of oscillation and then the contribution of the emission term $\mathbf{E}^{(-)}(\mathbf{r}, t)$ is entirely negligible. (We are assuming that the detector is at a relatively low temperature, as we have remarked in the preceding lecture.)

We must next sum the squared modulus of the amplitude (4.6) over all final states $|f>$ of the field, since no observations are ordinarily made of those states. One of the virtues of working with the expression (4.6) for the amplitude is that in the final state summation we can sum over <u>all</u> the states of a complete set; those final states which cannot be reached by the field for physical reasons are present in the sum but contribute nothing, either because the matrix elements leading to them vanish identically, or because the time integrals of the matrix elements vanish.

Thus the constraint represented by the conservation of energy, for example, is actually implicit in the structure of the time integrals in the sum of the squared amplitudes,

$$\sum_f |\langle a_f|U(t,t_0)|g\rangle|^2 = \left(\frac{e}{\hbar}\right)^2 \int_{t_0}^{t}\int_{t_0}^{t} dt'dt''\, e^{i\omega_{ag}(t''-t')} \sum_{\mu,\nu} M^*_{ag,\mu} M_{ag,\nu} \langle i|E^{(-)}_\mu(r,t') \times E^{(+)}_\nu(r,t'')|i\rangle, \tag{4.7}$$

which has been derived by using the relation

$$\langle f|E^{(+)}_\nu|i\rangle^* = \langle i|E^{(-)}_\nu|f\rangle$$

and the completeness relation $\Sigma_f |f\rangle\langle f| = 1$.

We have already discussed the need to average such expression as Eq. (4.7) over an ensemble of initial states $|i\rangle$ since the initial state is rarely known accurately in practice. We then find for the transition probability the expression

$$\begin{aligned} p_{g\to a}(t) &= \{\Sigma|\langle a_f|U(t,t_0)|g\rangle|^2\}_{\text{av over } i} \\ &= \left(\frac{e}{\hbar}\right)^2 \sum_{\mu,\nu} \int_{t_0}^{t}\int_{t_0}^{t} dt'\,dt''\, e^{i\omega_{ag}(t''-t')} \quad M^*_{ag,\mu} M_{ag,\nu} \\ &\qquad \mathrm{Tr}\,\{\rho E^{(-)}_\mu(r,t') E^{(+)}_\nu(r,t'')\} \\ &= \left(\frac{e}{\hbar}\right)^2 \sum_{\mu,\nu} \int_{t_0}^{t}\int_{t_0}^{t} dt'\,dt''\, e^{i\omega_{ag}(t''-t')} M^*_{ag,\mu} M_{ag,\nu}\, G^{(1)}_{\mu\nu}(rt', rt'') . \end{aligned} \tag{4.8}$$

The definitions of the density operator ρ of the field and of the first order correlation function $G^{(1)}$ have been given in the preceding lecture.

The foregoing discussion has assumed that the atom makes a transition to a specified final state $|a\rangle$. Counters employing discrete final states have received a certain amount of discussion recently. Bloembergen and Weber, for instance, have proposed using a scheme illustrated by Fig. 5.

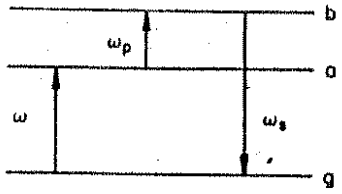

Figure 5

When the atom is excited to the state a by an incident field of frequency ω it is then raised to a higher level b by a pumping field at frequency ω_p. The emission of a photon with the sum frequency $\omega_s = \omega + \omega_p$ indicates the absorption of a photon from the incident field.

In the detectors used to date, however, the final states $|a\rangle$ of the atoms form an extremely dense set, or a continuum; the atoms are simply ionized, for instance. Since a counter of photoelectrons has only a limited ability to select among final atomic states (e.g., the counting of photoelectrons places only weak restrictions on their momenta), we have to sum the probability given by Eq. (4.8) over at least part of the continuum of final states $|a\rangle$. But not all ejected electrons can really be counted. Often they are ejected in directions for which the counter is insensitive or they are stopped by matter. The device might furthermore be built so as to introduce some explicit selection according to energies before detecting photoelectrons.

We shall not discuss the actual means used for detecting the photoelectrons in any detail here. Instead we shall assume simply that the probability that an electron ejected by photoabsorption is really registered is given by some function R(a). The way in which this function varies with the final state $|a\rangle$ of the electron-ion system will depend, in general, on the geometrical and physical properties of the actual counting device. If we now sum the probabilities given by Eq. (4.8) over the final states $|a\rangle$ using the probability R(a) as a weight, we find for the probability of detecting a photon absorption in our one-atom detector

$$\begin{aligned} p^{(1)}(t) &= \sum_a R(a)\, p_{g\to a}(t) \\ &= (e/\hbar)^2 \sum_{\mu,\nu} \int_{t_0}^{t}\int_{t_0}^{t} dt'\,dt'' \sum_a R(a)\, M_{ag,\mu} M^*_{ag,\nu} \\ &\quad \times e^{i\omega_{ag}(t''-t')}\, G^{(1)}_{\mu\nu}(rt', rt'') . \end{aligned} \tag{4.9}$$

We now separate the sum over the final states into two parts, a sum over the final electron energies and one over all other variables such as momentum directions, spin, etc. We do this by introducing the sensitivity function,

$$s_{\nu\mu}(\omega) = 2\pi\left(\frac{e}{\hbar}\right)^2 \sum_a R(a) M_{ag,\nu} M^*_{ag,\mu}\, \delta(\omega - \omega_{ag}) , \tag{4.10}$$

which contains contributions only from transitions with a fixed energy transfer, $\hbar\omega$. (Note that $s_{\nu\mu}(\omega)$, although it is written as a sum of delta functions, is actually a well-behaved function for the case we are considering since the sum over states $|a\rangle$ is really an integration over states with a continuum of energies.)

By making use of the sensitivity function and of the properties of the delta-function it contains we may write the counting probability in Eq. (4.9) in the form

$$p^{(1)}(t) = \frac{1}{2\pi} \int_{t_0}^{t} dt' \int_{t_0}^{t} dt'' \int_{-\infty}^{\infty} d\omega \sum_{\mu,\nu} s_{\nu\mu}(\omega)\, e^{i\omega(t''-t')}\, G^{(1)}_{\mu\nu}(rt', rt'') . \tag{4.11}$$

Since $s_{\nu\mu}(\omega) = 0$ for $\omega < 0$ we have extended the integral over the variable ω from $-\infty$ to $+\infty$. If we define the Fourier transform of the sensitivity function by

$$S_{\nu\mu}(t) = (1/2\pi) \int_{-\infty}^{\infty} s_{\nu\mu}(\omega)\, e^{i\omega t} d\omega , \tag{4.12}$$

we finally obtain

$$p^{(1)}(t) = \int_{t_0}^{t} dt' \int_{t_0}^{t} dt'' \sum_{\nu\mu} S_{\nu\mu}(t''-t')\, G^{(1)}_{\mu\nu}(rt', rt'') . \tag{4.13}$$

Eq. (4.13) represents the total transition probability when our shutter is open from time t_0 to t. To obtain the rate at which transitions occur we must differentiate with respect to t.

In general there is nothing very localizable in time about the absorption process. It is not possible to say that the photon has been absorbed in a particular interval of time small compared to the total period during which the shutter has been open. This becomes quite clear if we assume that the sensitivity $s_{\nu\mu}(\omega)$ is sharply peaked with a small width $\Delta\omega$. Then $S_{\nu\mu}(t'' - t')$ takes on nonvanishing values for $|t'' - t'| \leq 1/\Delta\omega$, which may be an arbitrarily long interval of time for small $\Delta\omega$. The degree of non-locality in time which enters the integral in Eq. (4.13) is, roughly speaking, just the reciprocal $1/\Delta\omega$ of the bandwidth of our device. If the bandwidth is narrow the counter measures an average of values of $G^{(1)}(rt', rt'')$ with t' quite different from t''. In optical experiments a narrow band sensitivity is usually reached by putting narrow band light filters in front of broad band counters, i.e., by "filtering" the correlation function $G^{(1)}$ rather than by discriminating between photoelectrons. Broad band counters are therefore, in this sense, somewhat more basic than narrow band ones.

In the limiting case of extremely broadband detection the detection process

becomes approximately local in time. We have already made some mention in the preceding lecture of an ideal photodetector. Such a detector, we shall assume, has a sensitivity function $s_{\nu\mu}(\omega)$ which is constant for all frequencies. To gain a quick insight into the meaning of this assumption we note that when the sensitivity function is a constant, $s_{\nu\mu}$, independent of frequency, Eq. (4.12) reduces to

$$S_{\nu\mu}(t) = s_{\nu\mu}\,\delta(t). \tag{4.14}$$

The photon absorption process then becomes, in effect, localized in time, and the transition probability given by Eq. (4.13) reduces to

$$p^{(1)}(t) = \sum_{\mu,\nu} s_{\nu\mu} \int_{t_0}^{t} G^{(1)}(rt', rt')\,dt'. \tag{4.15}$$

Now the assumption that $s_{\nu\mu}(\omega)$ is independent of frequency would be quite a difficult one to meet in practice for $\omega > 0$. When we take negative values of ω into account it becomes, strictly speaking, an impossible condition to meet since $s_{\nu\mu}(\omega) = 0$ for $\omega < 0$. But in fact neither of these troubles stands in the way of our constructing actual devices which approximate the behavior of ideal detectors arbitrarily well, as long as we agree to use them on radiation fields of restricted frequency bandwidth. Once we assume that the field excitations have finite bandwidth all we really require of our detector is that its sensitivity be constant over the excited frequency band. The detector then functions in an ideal way no matter how much the sensitivity varies outside the excited band.

To show that we need only be concerned to have the sensitivity remain constant over the band which is actually excited, we shall examine Eq. (4.11) for the transition probability in a little more detail. Let us begin by imagining that the time interval $t - t_0$ is exceedingly great, e.g., we let $t \to \infty$ and $t_0 \to -\infty$. Then if we let $K_{\mu\nu}(\omega)$ be the Fourier integral

$$K_{\mu\nu}(\omega) = \int_{-\infty}^{\infty} dt' \int_{-\infty}^{\infty} dt''\, e^{i\omega(t''-t')}\, G_{\mu\nu}^{(1)}(rt', rt''), \tag{4.16}$$

it is clear that $K_{\mu\nu}$ vanishes for frequencies ω lying outside the excited band. (e.g., The diagonal elements $K_{\mu\mu}(\omega)$ are simply proportional to the power spectra of the three field components.) We may then make use of $K_{\mu\nu}(\omega)$ to rewrite Eq. (4.11) as

$$p^{(1)}(t) = (1/2\pi) \int_{-\infty}^{\infty} \sum_{\mu\nu} s_{\nu\mu}(\omega)\, K_{\mu\nu}(\omega)\, d\omega \tag{4.17}$$

Now as long as $s_{\nu\mu}(\omega)$ takes on the constant value $s_{\nu\mu}$ over the excited band (and no matter how it behaves elsewhere) we may write Eq. (4.17) as

$$\begin{aligned} p^{(1)}(t) &= \sum_{\mu\nu} s_{\nu\mu} (1/2\pi) \int_{-\infty}^{\infty} K_{\mu\nu}(\omega)\, d\omega \\ &= \sum_{\mu\nu} s_{\nu\mu} \int_{-\infty}^{\infty} G_{\mu\nu}^{(1)}(rt', rt')\, dt', \end{aligned} \tag{4.18}$$

and the latter of these expressions again shows the locality in time of the photon absorption process which we noted earlier in Eq. (4.15), i.e., the two arguments of the correlation function in the integrand are the same.

In order to derive the foregoing result we imagined that the time interval $t - t_0$ was allowed to become infinite. To see the influence of the fact that the time interval has a finite length, let us define a time-dependent step function

$$\eta(t') = \begin{cases} 0 \text{ for } t' < t_0 \\ 1 \text{ for } t_0 < t' < t \\ 0 \text{ for } t' > t \end{cases}$$

Then the limits of the time integrations in Eq. (4.11), for example, may be extended from $-\infty$ to ∞ if we first multiply the correlation function in the integrand by $\eta(t')\,\eta(t'')$. This extension of the limits of the time integrations means that we may use once more an argument of the type which led to Eq. (4.18). But the difference is that the function $K_{\mu\nu}(\omega)$ must now be regarded as the Fourier transform

$$K_{\mu\nu}(\omega) = \int_{-\infty}^{\infty} dt' \int_{-\infty}^{\infty} dt''\, e^{i\omega(t''-t')}\, \eta(t')\, G_{\mu\nu}^{(1)}(rt', rt'')\, \eta(t''). \tag{4.19}$$

The bandwidth of this function will in general be different from that of the radiation present but the difference will only be significant if the period during which the shutter is open is extremely brief.

Let us suppose the bandwidth of the radiation present, i.e., of the function $G^{(1)}$, is $\delta\omega$. The bandwidth associated with the functions η is of order $(t - t_0)^{-1}$. The frequency width characteristic of $K_{\mu\nu}(\omega)$ is presumably of the magnitude of the larger of these two widths. Then if we assume that the sensitivity function of our detector only varies appreciably over an interval $\Delta\omega$, we shall secure an expression for the transition probability which reduces to the form of Eq. (4.15) as long as $\Delta\omega$ satisfies the two conditions

$$\Delta\omega \gg \delta\omega \qquad \text{and} \qquad \Delta\omega \gg (t - t_0)^{-1}.$$

The second of these conditions sets a lower bound $1/\Delta\omega$ to the length of time our shutter can be open if we want the behavior of our counter to remain ideal.

If we differentiate Eq. (4.15) with respect to time we find that the rate of increase of the transition probability, i.e., the counting rate, is

$$w^{(1)}(t) = \frac{dp^{(1)}(t)}{dt} = \sum_{\nu\mu} s_{\nu\mu} G_{\nu\mu}^{(1)}(rt, rt). \tag{4.20}$$

Having carried the tensor indices of the sensitivity and correlation functions far enough to illustrate their role in determining the transition probabilities we shall now eliminate them by imagining the field to possess a specified polarization $\hat{e}$. This can be accomplished in practice, of course, by putting a polarization filter in front of the counter. With the notation

$$\begin{aligned} E^{(-)}(r,t) &= \hat{e} \cdot E^{(-)}(r,t) \\ E^{(+)}(r,t) &= \hat{e}^* \cdot E^{(+)}(r,t) \\ G^{(1)}(rt, r't') &= \mathrm{Tr}\{\rho\, E^{(-)}(r,t)\, E^{(+)}(r',t')\} \\ s &= \sum_{\mu\nu} \hat{e}_\nu\, s_{\nu\mu}\, \hat{e}_\mu^*, \end{aligned} \tag{4.21}$$

Equation (4.20) may be rewritten as

$$w^{(1)}(t) = s\, G^{(1)}(rt, rt). \tag{4.22}$$

We have thus justified the assumption, made in the course of the simplified discussions given earlier, that an ideal photon counter can be constructed to respond, in effect, to the field at a given instant of time. Its counting rate is proportional to the first order correlation function evaluated at a single point and a single time.

In deriving the foregoing results we have employed the electric dipole approximation. The use of that approximation has been much more a matter of convenience than one of necessity. We could as well have retained the general coupling between the momentum of the atomic electrons and the vector potential. We would then have made use of correlation functions for the vector potential rather than for the electric field. The only difference in the calculations would then be a matter of