10

INTRODUCTION TO QUANTUM PHYSICS

Part VII of Physics - A New Introductory Course

Preliminary edition

Arthur K. Kerman

Leo Sartori

Edwin F. Taylor

QC174.1
.K39
1968

These notes are the third draft of this part of Physics, a New Introductory Course. Several additional chapters are being written, and the chapters included here are being revised again.

Copyright © 1967, 1968 by the Education Research Center
at the Massachusetts Institute of Technology

RECEIVED
FEB 26 1968
M. I. T. LIBRARIES

TABLE OF CONTENTS

0583935

Chapter 1. QUANTUM STATES

1. Introduction

During the 19th century physicists and chemists demonstrated convincingly that all matter consists of individual building blocks to which the ancient Greeks had given the name atoms. The modern workers deduced from indirect evidence that the dimensions of atoms are of the order of 10^{-8} centimeters.* They assumed that atoms and molecules obey the laws of classical mechanics and electromagnetism--laws that had been derived from the study of much larger objects. Some atomic phenomena could be explained fairly successfully on the basis of this assumption. For example, the kinetic theory, which treats each molecule as a small particle subject to the laws of classical mechanics, accounts for many properties of gases at low pressure.

As experimental technique improved and new phenomena were discovered and studied, serious difficulties in the classical description became apparent. The first strong signs of trouble appeared in the study of the radiation emitted by matter--both the line spectra of gases in a low-pressure discharge and the continuous "blackbody" spectrum radiated by hot bodies. The phenomena for which no adequate classical explanation can be given now encompass practically all branches of physics and chemistry. A partial list of such phenomena is given in Table 1.

Table 1. Some Properties and Phenomena Inadequately Explained by Classical Physics

In the field of atomic physics:	size of the atom stability of the atom discrete frequencies of light emitted by excited atoms effects on the atom of bombardment by light or by material particles
In the field of chemistry:	differences in chemical reactivity among different atoms strength of binding of atoms into molecules shapes and sizes of molecules
In the field of solid state physics:	differences in electrical and thermal properties between insulators, semiconductors, normal metals, and superconductors ferromagnetism specific heat of solids at low temperature cohesive energy of liquids and solids.

---

* For the methods of deducing dimensions of atoms see F. L. Friedman and L. Sartori, The Classical Atom, Addison-Wesley Publishing Company, 1965. Pages 2 and 46.

Table 1. (Continued)

In the field of nuclear physics:	dimensions of the nucleus discrete energies of photons and material particles emitted by some nuclei nuclear fission and fusion nuclear transformations under bombardment

The present notes are an introduction to quantum mechanics, the theory that successfully replaces the classical description for the phenomena of Table 1. We shall by no means provide an explanation for all the items that appear in the table. Our aim is to present the basic principles of quantum mechanics by analyzing a few key experiments. These experiments are not necessarily the ones that have been most important in the historical development of the theory. Some have been performed only recently; a few are quite idealized and might almost be described as thought-experiments. But this approach constitutes a reasonably direct path to the principal ideas. In later chapters, we shall illustrate some of the techniques of quantum mechanical calculation by applying them to problems of physical interest.

Some aspects of the quantum mechanical description of particle behavior appear at first to be quite strange. In order to interpret experimental results consistently one is obliged to accept some premises and conclusions that seem contrary to everyday experience and "common sense," and to abandon others that seem almost self-evident. But recall that the same things happen when one studies special relativity. In each case our everyday experience is based on observation of phenomena in a restricted domain--in one case the domain of low-velocity motion, in the other the domain of objects large compared to atomic dimensions. When we venture outside the boundaries of the familiar, we must be prepared to have our intuition fail.

Just as relativity reduces to Newtonian mechanics in the limit of low velocities, so does quantum physics reduce to classical physics in certain limiting cases. The behavior of macroscopic bodies generally (but not always) falls within the range of this "classical limit," although this is not so readily demonstrated as is the non-relativistic limit. The phenomena in which quantum mechanical effects are most apparent take place at the atomic or sub-atomic level.* If our everyday experience included working in an atomic or nuclear physics laboratory, we would constantly be confronted with the differences between the laws of quantum mechanics and those of classical physics, and the former would no longer be strange to us.

2. Identity and discreteness.

Most experiments on the atomic scale involve a large number of particles. To study electron-proton scattering, for example, it would not suffice to project a single electron at a single proton (even if this were feasible) and observe how the electron is deflected. Moreover, one

* But macroscopic size in itself is not sufficient to ensure that a phenomenon may be explained classically. Superconductivity and the operation of a transistor are but two common examples of the behavior of large objects that can be explained only by quantum physics.

cannot pick up the same electron and throw it at the same proton over and over, as one could in a similar experiment with macroscopic bodies. What one does, instead, is prepare a beam of electrons, aim the beam at a target that contains protons, and observe how many electrons are deflected through various angles. How can such an experiment tell us anything about the features of a single electron-proton encounter? The reason is that, as far as we can tell, every electron is identical to every other electron in all respects: each electron, has, within the accuracy of the best experiments, exactly the same mass, exactly the same charge, and so on, as every other electron. Protons are also identical to one another. Therefore if the beam has been carefully prepared, an experiment in which, say, 10^{10} electrons are hurled at the target and 10^6 of these are detected conveys information about the collision between one electron and one proton. The identity of atomic particles by class is the property that makes such an experiment possible.

In a strictly logical sense, the statement that all electrons are identical to one another is a statement true by definition. An electron is defined with a certain mass, charge, magnetic moment, and so forth. Therefore all electrons are necessarily identical to one another. If someone discovered two types of electrons which differed, say, in mass, we would give each a separate name, and the members of each newly-defined class would then be identical among themselves. Indeed, there exists a particle that is apparently identical to the electron in all properties except its mass, which is 207 times the mass of the electron. This "heavy electron" has a separate name: the negative mu meson.* The existence of mu mesons causes no difficulty because they can be readily distinguished from ordinary electrons. It would be a serious problem if electrons came with a continuous range of mass values, because such particles would be hard to distinguish from one another. Fortunately, the masses and charges of atomic and nuclear particles take on only discrete values, separated by appreciable intervals. In the next section we shall see that the energies of atomic and nuclear systems likewise take on only certain discrete values. The same is true of angular momentum. We say that charge, mass, angular momentum, and (sometimes) energy are quantized. Quantum mechanics, as the name implies, is built upon this central feature of nature, which makes it possible to sort particles into classes and to study the members of each class as identical particles.

The property of identity is not, of course, limited to atomic-sized particles. One can conceive of two identical grains of sand: sand is (largely) composed of silicon dioxide, and all molecules of silicon dioxide have the same composition. Nevertheless, the chance of finding two strictly identical grains of sand is negligible. The masses of two grains can differ by as little as one molecular mass. This difference is so small compared with the mass of a grain that mass values, for all practical purposes, form a continuum. When differences other than mass

---

* A similar example is provided by the neutrino. For thirty years it was assumed that all neutrinos are identical. Then in 1964 an experiment was performed which allowed one kind of neutrino (emitted in the decay of some radioactive nuclei, for instance cesium 134-- see the next section) to be distinguished from another kind (emitted in the decay of the mu meson). Now we define two kinds of neutrinos!

values are taken into account, the number of possible variations becomes larger than the number of grains of sand on all the world's beaches! This example illustrates the fact that the concept of identity has limited usefulness in the physical study of macroscopic bodies.

3. Energy Levels

Atomic energy levels. It is difficult to perceive atoms individually. Only because of the identity among atoms of a given kind is it possible to study their properties. Test tube chemistry depends upon the assumption that many identical processes take place concurrently in the tube, a number of processes sufficient to yield detectable results. The same assumption lies behind the study of atomic spectra, a study which led to most of the early developments of quantum mechanics. Let a small drop of mercury be heated in an evacuated container until the mercury vaporizes. Individual atoms of mercury which fill the container are separated, on the average, by distances equal to thousands of times their diameter. Therefore they do not interact with one another significantly. Because they are identical they respond in a similar manner to external influences. In particular, an electric discharge through the vapor produces light whose spectrum is characteristic of the individual mercury atoms (although the intensity depends on the number of atoms present).

The spectrum of a luminous gas may be measured by means of a diffraction grating (figure 1).

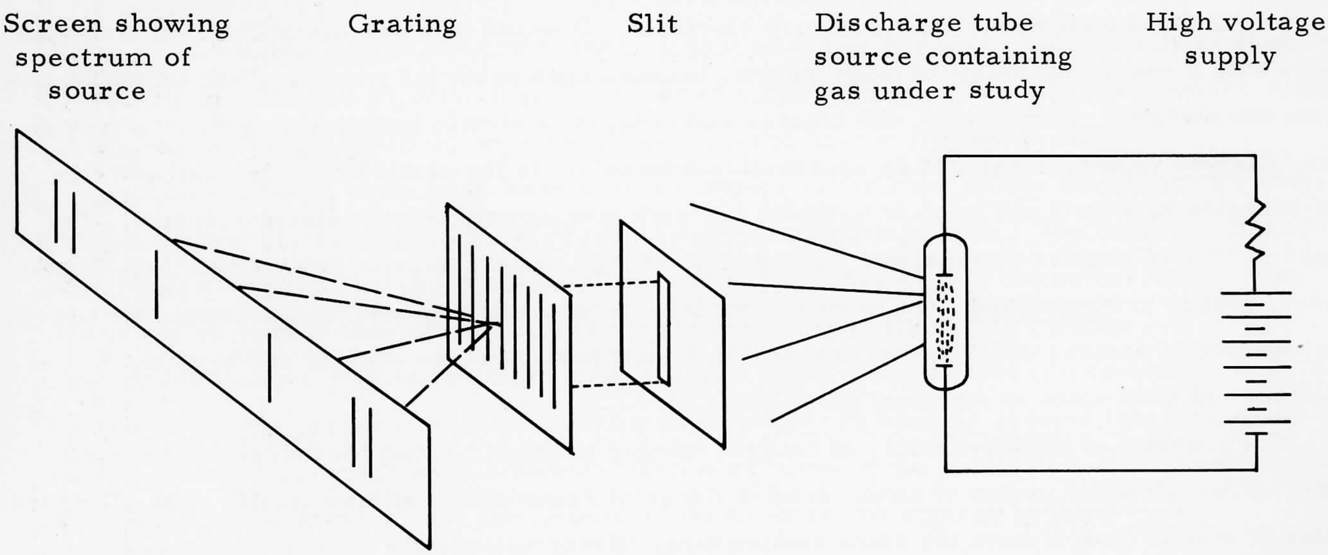


Figure 1. Schematic diagram of an experiment by which a spectrum of a vaporized sample may be observed. The intensity plot of Figure 2 is obtained by measuring the darkening on a photographic plate substituted for the screen.

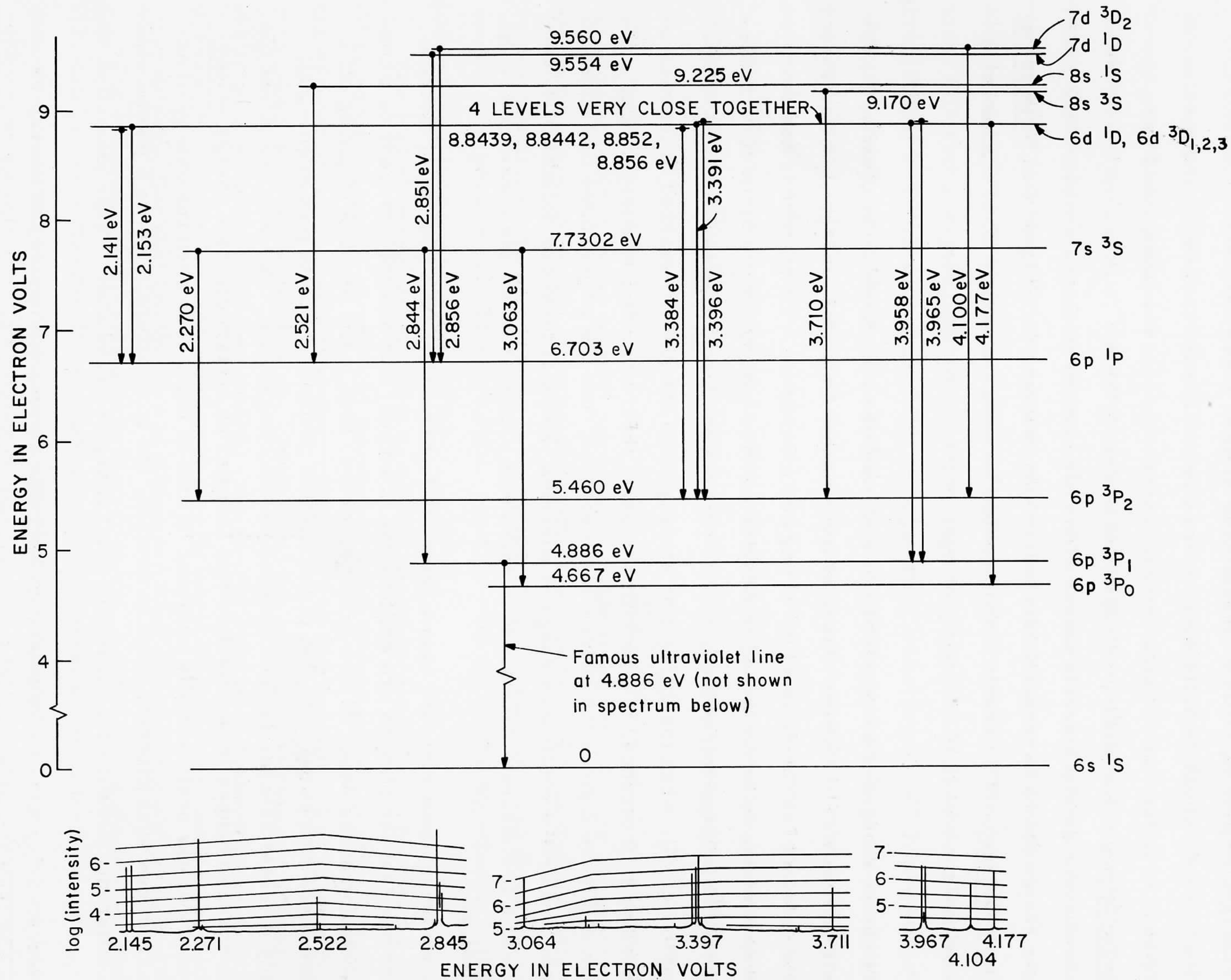


Figure 2. Part of the atomic spectrum of mercury. The spectral peaks correspond to transitions between energy levels shown in the upper diagram. Labels on the energy levels (on the right) are those conventionally used to describe the levels in question. The entire atomic spectrum of mercury has many more peaks than those in the limited energy range shown here; and there are many more energy levels of mercury than those shown in the upper diagram. Spectrum taken from the American Institute of Physics Handbook, Second Edition, McGraw-Hill Book Company, 1963.

The spectrum of mercury vapor is shown in Figure 2.* The characteristic feature of this spectrum is the presence of sharp peaks--high intensity over narrow ranges of frequency. The spectra of discharges containing other elements show similar peaks, at frequencies characteristic of the atom. These characteristic spectra are commonly used to identify elements in the laboratory, as well as in the sun and other stars. Atomic spectra are interpreted using the concept of energy levels of the atoms, as explained in what follows.

In order to account for the discrete nature of atomic spectra we need to preview one result of Chapter 2--a result you have probably met before (for instance, in Chapter 8 of Physics, a New Introductory Course). This result is the relation between the energy E of a quantum of light--a photon--and the classical frequency of light ν:

$$E = h\nu \tag{1}$$

A beam of light can be thought of as consisting of a stream of such photons. In equation (1) the constant of proportionality h between classical frequency and quantum energy is called Planck's constant. This constant has the dimensions of angular momentum. Its magnitude has been measured in experiments to be described in Chapter 2; it is approximately 6.6×10^{-34} joule-second or 4×10^{-15} electron-volt second.

According to Eq. (1), when the diffraction grating sorts the photons emitted in the mercury discharge according to frequency, it automatically sorts them according to energy. The scale on the abscissa of Figure 2 gives the energy of the photons; thus the graph shows that although photons are emitted over a continuous range of energies, some of them are definitely grouped in energy near a few discrete values. It is with these photons that we are concerned in the present discussion.

What can we learn about mercury atoms from the fact that they emit photons of discrete energies? Assume that the photons are emitted one at a time; this assumption leads to a consistent interpretation of the data. If energy is conserved, then in the process of giving off a photon the atom loses energy equal to that of the emitted photon. Discrete photon energies therefore imply discrete changes in the energy of the emitting atoms. One way to account for these discrete energy changes is to postulate that there exist certain discrete energy values which the mercury atoms preferentially possess. If this is true, then when the energy of an atom changes from one such discrete value to another, the atom naturally emits a photon whose energy is the difference between the initial and final energies of the atom. This idea--that there

---

* Four prominent lines in the visible region of the mercury spectrum may be observed using a plastic diffraction grating that costs a few pennies. A fluorescent lamp exhibits the blue line (2.85 eV), the green line (2.27 eV), and the yellow line (2.15 eV) shown in Figure 2; and in addition a red line not shown in the figure. A mercury vapor street lamp--identifiable by its bluish color--yields a spectrum free of the distracting background rainbow from the fluorescent coating of the fluorescent tube.

exist discrete energy values for atoms--is due primarily to Niels Bohr; it was one of the major steps in the development of quantum theory.

The discrete energies which an atom may have are called the energy levels of that atom. At ordinary temperatures most of the atoms in a given gas sample are in the level of lowest energy, called the ground level. Heat or electric discharge, both of which produce energetic collisions between atoms, can be used to transfer some of the atoms to levels of higher energy, called "excited levels." In dropping back subsequently from a higher energy level to a lower energy level, the atom emits the energy difference in the form of a photon. The hypothesis of atomic energy levels receives strong support from the pattern of emitted photon energies. The many different energies carried by photons emitted from one kind of atom can be interpreted consistently in terms of energy differences among a relatively small number of energy levels of the atom. Figure 2 shows a somewhat simplified energy level diagram for mercury. Photon energies in the spectrum of Figure 2 can be seen to correspond to energy differences between pairs of energy levels.

Notice that, according to the hypothesis of discrete energy levels, the observed discrete frequencies characteristic of a given element should satisfy relationships of the following sort:

$$\nu_1 + \nu_2 = \nu_3 + \nu_4 \tag{2}$$

Many relationships of the type (2) can indeed be discerned from Figure 2. Historically, the argument proceeded in the opposite direction: the existence of relationships such as (2) was noticed empirically by Ritz (1908), Rydberg (1900) and others, and was given the name "Ritz combination principle." This principle, together with the photon hypothesis, led Bohr to the idea of discrete energy levels.

The idea of energy levels permits a consistent interpretation of the spectra of all atoms. After several decades of time and the work of many physicists and physical chemists, the spectra of most of the elements have been interpreted in this manner. The central result of this continuing work is the same as the result found for mercury: the energy of every photon emitted from a given atom can be equated to the energy difference between some pair of energy levels in the atom. Moreover, the number of energy levels required for this interpretation is considerably smaller than the number of different photon energies for which the levels account.

The existence of discrete atomic energy levels, inferred from the observation of line spectra, can be more directly demonstrated by using a beam of electrons as a probe. This was first done by Franck and Hertz in a classic experiment performed in 1914. Figure 3 is a schematic sketch of their set-up. Electrons are emitted by the filament F which is at the negative potential -V. In front of the anode A is a metal grid G which is maintained at a small positive potential (about .5 volts). The entire apparatus is enclosed in a container filled with mercury vapor. The current collected by the anode is measured as a function of the variable voltage V.

What sort of results does one anticipate for this experiment? Presumably, a given electron

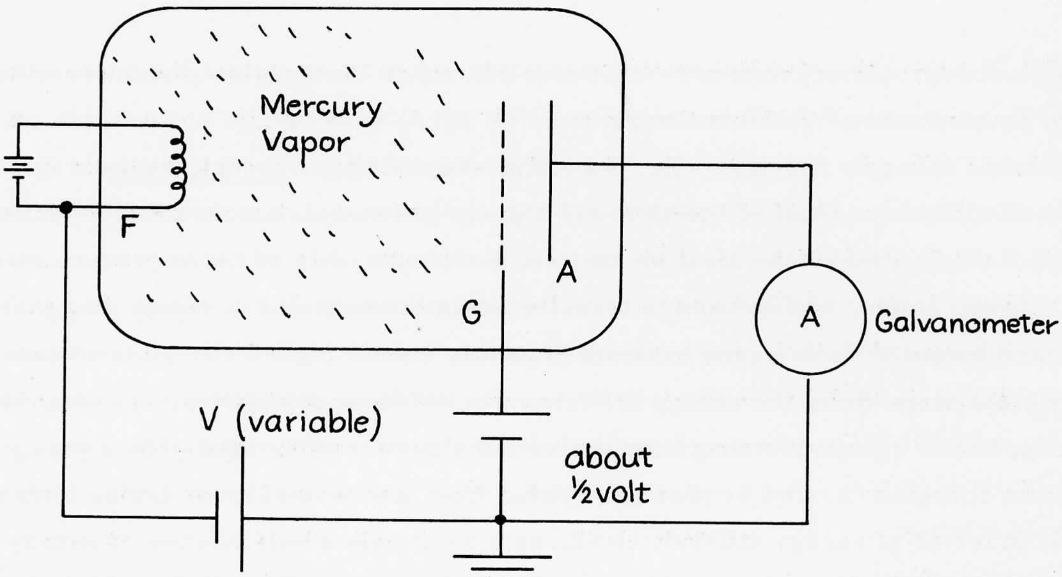


Figure 3. Experimental arrangement for Franck-Hertz Experiment

will be collected at the anode if it reaches the grid with a kinetic energy greater than .5 eV. Thus for any value of V, every electron emitted at the filament ought to be collected, unless it loses enough energy in collisions with mercury atoms to fall below the half-eV barrier. Notice that in an elastic collision with an atom, an electron can lose only a minute amount of energy.* A very large number of such collisions would have to take place before the electron suffered a significant energy loss. Since the number of electrons emitted at the filament is a fairly rapidly increasing function of its potential,** we expect that the anode current will increase steadily with increasing V.

Figure 4 shows the results obtained by Franck and Hertz. At first, the anode current in-

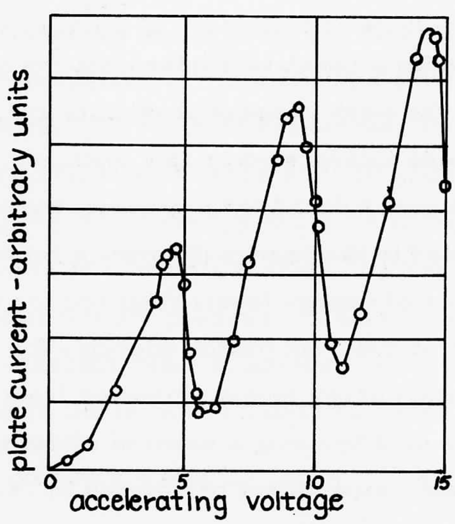


Figure 4. Results of the Franck-Hertz Experiment

---

* $(\Delta KE)_{max} = (KE)_0(4m/M)$, where m is the mass of an electron and M the mass of the atom.

** This can be verified by performing the same experiment in an evacuated chamber. For low voltages, the relation between current and voltage is approximately $I \propto V^{3/2}$.

creases about as expected. But just under 5 volts accelerating potential, the current drops sharply. This is a strong indication that many electrons are losing just about 5 eV. in a single collision with an atom. Such a collision is necessarily inelastic; the energy cannot go into kinetic energy of motion of the mercury atom as a whole. Therefore, it must go into internal energy of the atom.* We conclude that the atom has an "excited level" at an energy about 5 eV. This conclusion is bolstered by the behavior of the current above $V = 5$ volts. For additional increases in voltage, the current once more increases, indicating that the electrons are losing 5 eV and not more in their inelastic collisions. This confirms that the energy level is indeed discrete. At about twice the voltage of the first drop, another sharp drop takes place. This is interpreted as showing that many electrons have suffered two inelastic collisions, losing an equal amount of energy in each. Several more such drops were observed, all separated by equal intervals of accelerating potential. From the best mean value for this interval Franck and Hertz concluded that the energy of the excited level is 4.9 ± 0.1 eV.;** this value is in quite good agreement with the value 4.886 eV. obtained from spectroscopic evidence. (See Figure 2.)

The Franck-Hertz experiment had a great impact on the development of quantum theory, since it indicated quite dramatically the existence of a discrete energy level. The original procedure, as described here, does not provide a useful tool for studying in detail the energy levels of atoms; it detects only the first excited level.+ But the method was soon generalized to overcome this restriction, and has provided much useful information.

Nuclear energy levels. Discrete energy levels exist not only in atoms but also in molecules and in nuclei. Figure 5 shows a "nuclear spectrum" of gamma rays (high energy photons) given off by a sample of barium 134 nuclei created in the decay of Cesium 134. The energies of these photons are some 10^5 times greater than those of the photons emitted in atomic transitions. Nevertheless one observes a discrete spectrum of peaks, just as in the atomic case. The existence of peaks in both cases suggests a similar theoretical interpretation in terms of energy levels. Just as atomic spectra can be described in terms of atomic energy levels, so the spectrum of Figure 5 can be made to fit a consistent scheme of nuclear energy levels. Such a scheme for barium 134 is indicated in the upper part of Figure 5. Similar energy level dia-

---

* The energy might be used to ionize the atom; but this possibility can be excluded by a number of arguments. (See the problems.)

** The value of V at the first drop is not the best estimate for the energy of the excited level, because we don't know the energy distribution with which the electrons are emitted from the filament.

\+ This statement is not precisely true. As Figure 2 shows, mercury does have a level at a slightly lower energy, about 4.7 eV. This happens to be a so-called "meta-stable state;" transitions between this level and the ground level (and vice versa) are very improbable. Notice that no emission line corresponding to this transition appears in the figure.

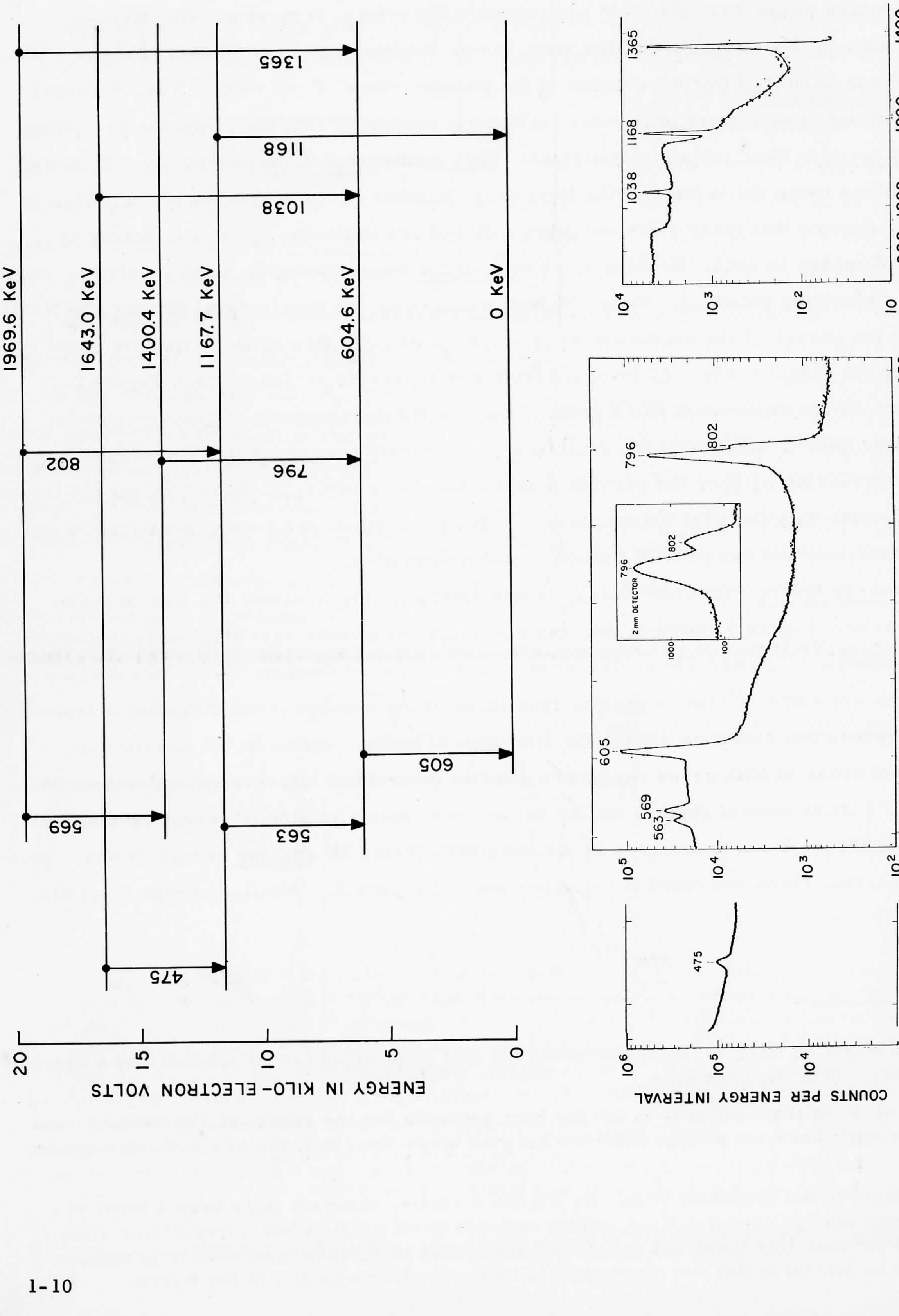


Figure 5. Nuclear spectrum of barium 134. These spectral peaks correspond to transitions between energy levels shown in the upper diagram. Taken from R. A. Brown and G. T. Ewan, Nuclear Physics, 68, 325 (1965).

grams have been derived for every stable nucleus.*

Here are some details of the experiment which yields the nuclear spectrum shown in Figure 5. Cesium 134 is a radioactive nucleus that decays to barium 134 by emitting a beta-particle (a high-speed electron) and a neutrino. The resulting barium nucleus is left in a level of high energy--one of the top three levels shown in Figure 3. The excited nucleus drops to lower energy levels, emitting gamma rays in the process.

How can one measure the energy of a high-energy photon? Low-energy photons emitted by mercury atoms (Figure 2) were measured with a simple diffraction grating. We understand the classical theory of diffraction gratings--the way in which they are used to measure wavelength and therefore, indirectly, frequency. The step from classical to quantum physics can then take place through the relation $E = h\nu$. Structures analogous to diffraction gratings can also be used to measure energy of high energy photons.** The structures are crystals in which the atoms or molecules form the grating. However, a simpler, easier, and more efficient method makes use of the quantum processes which take place when high-energy photons strike the semi-conductor germanium. These quantum processes are described pictorially and in a somewhat simplified fashion in Box 1. Perhaps it is unfair to use the quantum properties of a solid to demonstrate the quantum properties of nuclei. But experiments using germanium detectors give results identical to experiments on the same nuclei using the classically-understood diffraction gratings.

Box 1. Detection of Gamma Rays with a Germanium Detector

High energy photons (gamma rays) to be detected are allowed to fall on a specially-prepared germanium detector pictured in Figure 6. In passing through the germanium a photon can in-

---

* The energy levels of nuclei can also be investigated using other particles as probes. Let a beam of high energy protons fall on a target and measure the energies of the scattered protons. Some will be observed to have lost discrete amounts of energy as a result of inelastic collisions. One infers that these discrete energies have been transferred to the nuclei in the target, causing them to change to levels of higher energy. This proton-nucleus experiment is similar in principle to the Franck-Hertz electron-atom experiment discussed above. From the changes in energy of the emerging protons one can derive the nuclear energy levels. See PSSC Advanced Topics A-6 on Atoms, Molecules, and Nuclei, D. C. Heath and Company, Boston, Mass. (This footnote will be expanded into text material in the next revision of these notes.)

** J. W. M. DuMond, Review of Scientific Instruments, 18, 626 (1947).

teract with it in one of two ways:

(1) The photon can be entirely absorbed by an atom with the ejection of an electron with very high energy as shown in path a of Figure 6 ("photoelectric-effect"--see Section 1 of Chapter 2). This electron then moves rapidly through the crystal, tearing other electrons loose from atoms in its path. The many electrons thus released drift through the crystal under the influence of an applied electric field ("electron conduction") and are collected at one electrode. In addition the "hole" left behind when a given electron is torn from an atom can be filled by an electron from a neighboring atom, leaving a hole, which in turn can be filled from another nearby atom, and so on. By multiple steps of this kind the "holes" propagate in a direction opposite to the freed electrons in the applied electric field ("hole conduction") and are collected at another electrode. Now the number of electron-hole pairs created as a result of the total absorption of the incident photon is proportional--it turns out--to the initial energy of the photon. These electrons and holes are received in a pulse at the electrodes. The total pulse of charge collected by the electrodes after an event is therefore also proportional to the initial energy of the incident photon. The horizontal energy scale of Figure 5 is derived (after appropriate calibration) from the size of pulses received at the electrodes. The "peaks" of intensity correspond to the discrete energies of the incident photons.

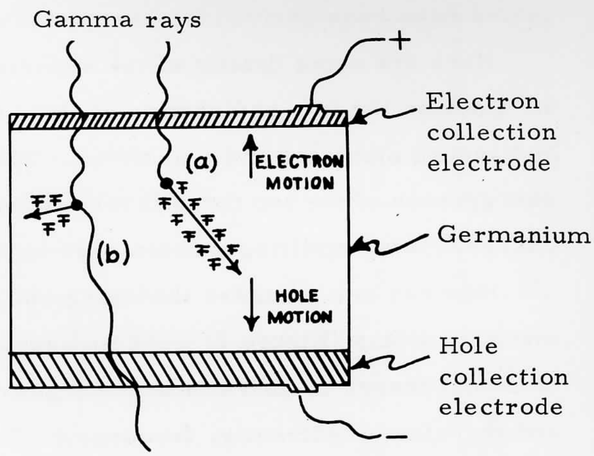


Figure 6. Schematic diagram of a germanium detector for high energy photons (gamma rays). Adapted from G. T. Ewan and A. J. Tavendale, Canadian Journal of Physics, 42, 2286 (1964).

(2) A second way for the photon to interact with the crystal detector is to scatter elastically from an electron ("Compton scattering"--see Section 1 of Chapter 2). The scattering electron receives part of the energy of the incident photon, while the remainder of the energy is carried out of the crystal by the scattered photon (path b of Figure 6). The electron creates electron-hole pairs in the crystal just as do the photoelectrons in case (1). However, in the present case only part of the energy of the incident photon is transferred to electron-hole pairs. Therefore the total charge collected at the electrodes is less than the amount which corresponds to the entire energy of the incident photon. Moreover, charges collected for many such events will not be grouped around discrete values of energy but will be spread all over the spectrum. This absorption of only part of the incident photon energy by the detector in process 2 is responsible for the high "background count" in Figure 5. The background count is concentrated at low energies, since every such count corresponds to less energy than the initial energy of the incident photon responsible for the count.

4. Quantum States

We emphasized in Section 2 the importance of the identity of particles in the development of quantum mechanics. The existence of discrete energy levels, discussed in Section 3, makes it possible to refine the concept of identity. Even though all mercury atoms are made up of the same constituents,* we can surely distinguish mercury atoms in their lowest energy level from mercury atoms in the first excited level. In this sense the two kinds of mercury atoms are not identical. Therefore we classify mercury atoms according to their energy levels. Is this as far as the class of mercury atoms can be divided? Perhaps the sub-class "mercury atoms in their lowest energy level" can be further subdivided according to some new experimental criterion. Are we condemned to analyze characteristic after characteristic according to which different mercury atoms can be distinguished? No purely philosophical answer can be given to this question; the issue is how natural objects behave. For the particular case of mercury atoms in their lowest energy level no further subdivision has been found: mercury atoms in their lowest energy level are, as far as we know, completely indistinguishable from one another.

Mercury atoms are only an example. All atomic and sub-atomic systems that have been investigated share with mercury atoms this fundamental character, that classes can be found which apparently cannot be subdivided any further. The members of each such class are completely indistinguishable from one another. We call such a class a quantum state.

Let a class of identical particles or systems be sufficiently described to distinguish the class from all other classes. Then we say the particles or systems are in the quantum state specified by that description. This definition of a quantum state is deeply rooted in the nature of atomic and sub-atomic systems and in our experience in investigating them. The full significance of the definition and some of its subtle implications will become clear only as we proceed with this study.

How does one know when a sufficient description has been given to define unambiguously a quantum state for a given system? How can we be sure that some crucial characteristic--as yet undetected by experiment--has not been left out of the description? We cannot be sure. The historical development of quantum mechanics has been in large part the search for observables that define quantum states and for values that these observables take for particular states. New classifications are sometimes found by the combination of an educated guess and a confirming experiment. The discovery that there are two kinds of neutrinos (see footnote on page 3) is an example of this process. We shall see later (Chapter 13) that angular momentum is an important quantity in specifying some atomic states. Angular momentum is, like energy, quantized (i.e. restricted to definite discrete values); the particular values of angular momentum are specified by "angular momentum quantum numbers." Some classifications required to define quantum

---

* Mercury has several isotopes, which differ according to the number of neutrons in the nucleus. We consider here a single one of these isotopes.

states are generalizations of classical concepts, as angular momentum is; others involve new, strictly quantum mechanical concepts. Occasionally, properties thought to be part of the description of a quantum state turn out to be irrelevant or without meaning. Sometimes one or more properties of a quantum state are not discrete but can take on any value in a continuous range (example: the kinetic energy of a free particle--Chapter 9). More than one alternative method may sometimes be used to specify a single quantum state. For example, the labels at the right side of the mercury energy level diagram--Figure 2--describe the states by a shorthand that specifies the various angular momenta (as well as one additional quantum number called the "principal quantum number"). The notation $6p\ ^3P_0$ defines, in terms of principal quantum number and angular momenta, the same state as does the value 4.667 eV in terms of energy. In summary, one cannot be sure in advance what properties are involved--and what ones are excluded--in a description sufficient to determine quantum states unambiguously. For a given system these classifications are obtained by guess, by intuition, and by inspiration; from formalisms and theories and symmetry arguments. No matter how the classifications originate, they are all tested ultimately in the arena of experience.

Quantum mechanics is the study of quantum states. In developing quantum mechanics we not only try to discover what properties appear in the specification of quantum states. We also study changes of systems from one quantum state to another. Some changes do not take place: electrons do not change into protons. Other changes do take place: an atom drops from a higher energy level to a lower energy level, emitting a photon; a neutron becomes ("decays into") an electron plus a proton plus a neutrino. The study of these transitions from one state to another--the nature of the transitions and their probability of occurrence--is another central preoccupation of quantum mechancs.

5. Photons

In the next few chapters we begin our study of quantum states by analyzing the polarization states of photons. For several reasons photons provide an attractive first example. Photons in the visible part of the spectrum can be studied especially easily. They are easy to produce, easy to experiment with (with simple mirrors, lenses, polarizers, and so on), and easy to detect. The human eye is a photon detector with an enormous range of adaptation. On a sunny day trillions of photons may strike the eye each second without causing either "saturation" or discomfort, yet on a dark night a man may be able to see a star so faint that only a few hundred photons enter his eye each second. In carefully arranged experiments a human eye can detect as few as 5 photons. The photomultiplier tube is a photon detector with smaller range of adaptation than the human eye but with more dependable sensitivity at low levels of illumination. A photomultiplier tube can detect photons of visible light with an efficiency of a few percent: a count will be recorded for every 5 to 50 photons which fall on its photosensitive surface.

Under ordinary circumstances photons do not interact with one another. One beam of light (one stream of photons) can pass through another beam without any effect on either. In a single beam the behavior of photons is independent of the intensity of the beam; the photons behave (most conveniently!) in a manner independent of one another. This property of photons allows a great simplification in instrumentation: We carry out experiments with a beam of visible light intense enough to be easily detected by the unaided eye. We are assured that the same results will be obtained, on the average, in experiments with small numbers of photons. (Under extraordinary circumstances--involving high intensity light in the presence of matter, or photons of very high energy--advanced theory predicts that photons will interact with one another. Such an interaction has not yet been observed. In any event the moderate-intensity beams of visible light described in early chapters of this book exhibit no interaction.)

The quantum states of photons in a beam can be specified with comparative ease, as we shall see in Chapter 3. One needs to give only the direction of the beam, the energy of photons in the beam, and their polarization. We shall concentrate on the polarization properties. Photons may be linearly, circularly, or elliptically polarized. The polarization can be studied by means of simple optical devices in experiments that can be carried out on a tabletop. These experiments have a straightforward interpretation in terms of the classical wave theory of light. But when we try to understand them on the basis of the photon picture, we are led to conclusions that form the basis for quantum mechanics. Thus, the interpretation of polarization phenomena in terms of photon states brings insight which will help us to understand the behavior of other systems, for which no classical explanation whatever can be given.

Photons have one major disadvantage as a vehicle for our preview of quantum mechanics: being intrinsically relativistic particles, they are readily created and destroyed. A process may begin with one photon present and end with two, or with none. A satisfactory description of the interaction of photons with matter requires a more advanced version of the theory, called quantum electrodynamics. In our present introductory treatment we will not be able to discuss details of the emission, absorption, and scattering of photons by atoms and molecules. Instead we will confine ourselves to a description of the quantum states of photons. The devices used in this study (polarizers, diffraction gratings, calcite crystals, etc.) will be characterized by their net effect on these photon states, with no present concern for how these effects come about. By proceeding in this way we can use photons to present quantum theory in a simple way, while avoiding the complications that must be faced in a complete dynamical treatment.

In brief our program will be as follows: Suppose that one takes seriously the existence of the photon as something analogous to a particle. Then a few simple table top experiments with polarized light considered as a photon beam will exemplify some of the major physical and mathematical features of quantum physics. With appropriate modification these same results will then be found to apply simply, naturally--and correctly!--to beams of other particles and then to systems of one or more particles. In order to begin this program we recall, in the following chapter, the evidence for the model of light as beams of photon-particles, and review the fundamental properties the model assigns to these particles.

CHAPTER 2. PROPERTIES OF PHOTONS

1. Evidence for the existence of photons.

Observe a dimly-lit wall, such as that of a darkened room at night with a small amount of light entering through windows or under a door. The wall looks neither dark nor light, but mottled, with a granular appearance that changes unpredictably with time. This granularity results from the granularity of light itself, composed of individual granules called photons.

Under the circumstances just described one does not actually observe individual photons. Rather, one is observing statistical fluctuations about some average illumination. For example, from a given square centimeter of the wall, 1000 photons, then 950, then 1030 photons may enter the eye of an observer in three consecutive 0.1 second intervals. It is these fluctuations--rather than the observation of individual bursts of illumination--which give visual evidence for the existence of photons. At higher levels of illumination, the fractional fluctuations are so small that they are not noticed. (For a discussion of fluctuations, see Physics, a New Introductory Course Chapter 9.)

Fluctuations in the illumination of a dimly-lit wall are correctly explained in terms of photons*. But the human eye is notoriously undependable as a quantitative instrument at low levels of illumination. Therefore one hesitates to base a belief in the existence of photons upon visual observations alone. And of course, such observations tell us nothing whatever about the energy of a photon. More convincing are experiments whose quantitative interpretation requires the existence of photons.

The fundamental formula for the photon description of light is that which relates the classical frequency ν of a light beam and the energy E of individual photons which compose the beam:

$$E = h\nu \tag{1}$$

This expression has already been used, in Chapter 1, to infer energy differences between atomic energy levels from the frequencies of observed line spectra. The constant of proportionality h--Planck's constant--between classical frequency and quantum energy is central in more than the study of photons. The very existence of Planck's constant (i.e., the fact that it is different from zero) accounts for the stability of atoms against collapse, as we shall see in a later chapter.

The photoelectric effect. Equation (1) was first postulated in 1900 by Max Planck in order to account for the spectrum of the radiation ("blackbody radiation") emitted from a small hole in a furnace. In studying the photons present in a furnace or other hot cavity, one is concerned with the thermal equilibrium that exists between photons of all possible energies and the walls of the furnace. We shall not discuss here the thermodynamic arguments used in such a study. Instead, we turn to a more direct testing ground for the photon hypothesis: the photoelectric

* Albert Rose, "Quantum Effects in Human Vision" Advances in Biological and Medical Physics, Vol. 5, New York, Academic Press, 1957.

effect; the ejection of electrons from a metallic surface when irradiated by ultraviolet (and sometimes by visible) light. The phenomenon was first observed by Hertz in 1887, and was subsequently studied intensely by Lenard and others. Its explanation was given by Einstein in 1905, when little besides the mere existence of the effect was known. Einstein boldly asserted that Eq. (1), which Planck had proposed in a quite limited context, is actually a universal property of electromagnetic radiation. Einstein's explanation of the photoelectric effect, based on this hypothesis, was fully confirmed when the complete quantitative nature of the effect was established some ten years later. Not only was Eq. (1) verified, but an accurate value for the constant h was obtained.

Assume that a beam of light of frequency ν consists of photons of energy $h\nu$, as Einstein asserted. Let such a beam fall on a metal target and suppose that a single electron in the target absorbs one of the photons in the beam. The resulting increase in the energy of the electron may be sufficient to enable it to leave the metal. An electron driven from a surface by light in this manner is called a <u>photoelectron.</u> The kinetic energy of the photoelectron can be written

$$\frac{1}{2} mv^2 = h\nu - W \qquad (2)$$

In this expression, W represents the energy which the electron loses on its way out, less the energy it originally possessed. Specifically, we can write

$$W = \Delta E_{coll} + W_1 - \frac{1}{2} mv_o^2 \qquad (3)$$

where $\frac{1}{2} mv_o^2$ is the electron's initial kinetic energy and ΔE_{coll} is the energy it loses through collision with other electrons or with atoms while moving toward the surface. W_1 represents the <u>loss</u> in kinetic energy which accompanies the potential energy <u>gained</u> by the electron while crossing the surface against the attraction of the atoms left behind. Notice that whereas W_1 is presumably a property of the surface and is the same for all electrons, the other two terms in Eq. (3) vary from one electron to another, and therefore so does W. Consequently, according to Einstein's hypothesis, the photoelectrons which leave the cathode have a distribution of kinetic energies. The fastest possible photoelectron is one which has the greatest initial kinetic energy (call this maximum initial energy $\mathcal{E}$), and which manages to escape with no collision loss. For such an electron W has its minimum possible value W_{min}:

$$W_{min} = W_1 - \mathcal{E} \equiv \phi \qquad (4)$$

The quantity ϕ, defined by Eq. (4), is characteristic of a material, and is called the <u>work function.</u> The maximum kinetic energy of the photoelectrons can now be written as

$$\frac{1}{2} mv_{max}^2 = h\nu - \phi \qquad (5)$$

The kinetic energy of an electron can be measured by determining what retarding potential it can overcome. The first careful experiments of this type on photoelectrons were carried out by Millikan in 1916: we describe such an experiment and show how Einstein's relation (2) pro-

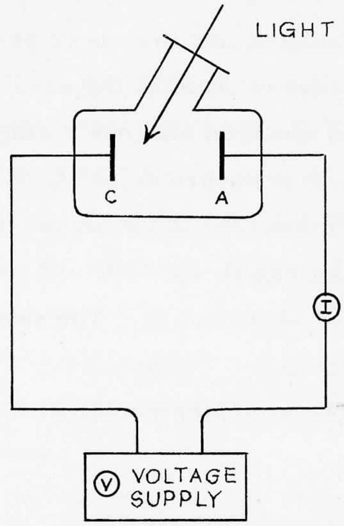


Fig. 1. Schematic diagram of the photoelectric experiment.

vides a complete explanation whereas classical wave theory fails. Figure 1 is a schematic diagram of the apparatus. Monochromatic light falls on a clean sodium cathode C ejecting photoelectrons. The anode A is maintained at a potential V relative to C; V can be either an accelerating or decelerating voltage. The experiment is performed in high vacuum so as to minimize the loss of energy by the photoelectrons due to collisions with gas molecules. In each run the wavelength and intensity of the incident light are kept constant, and the current collected by the anode is plotted as a function of the voltage V. Figure 2 shows Millikan's results. For wavelengths greater than 6800 Å, no current could be detected, even with a very strong accelerating potential, no matter how intense the incident light; below this critical or threshold wavelength, the measured current increased as the retarding potential was decreased, and

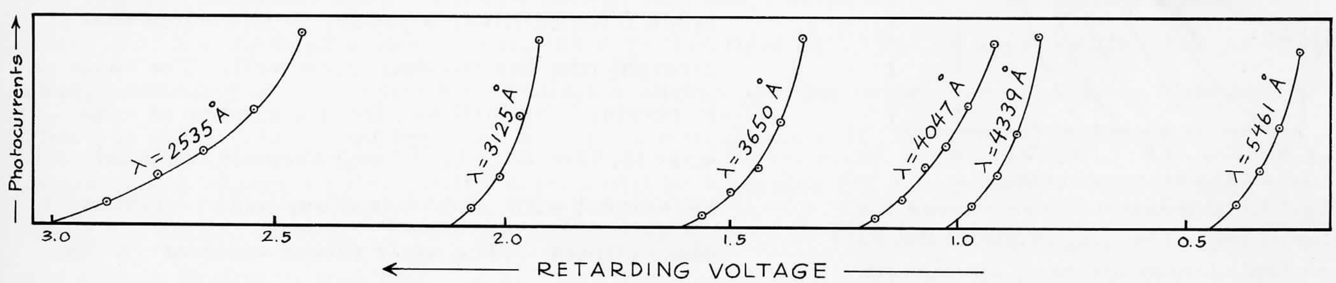


Fig. 2. Millikan's photoelectric experiment with a sodium cathode. Current vs. retarding potential for light of several wavelengths. (Retarding potentials have been corrected for contact potential difference between anode and cathode.)

eventually reached a saturation value. (See Figure 3, which is taken from a similar experiment with aluminum carried out by Compton and Richardson.)

On the basis of classical electromagnetism, one would attempt to explain the photoelectric effect roughly as follows: the oscillating electric field associated with the incident light wave "shakes" the electrons in the target and sets them into sympathetic oscillation. As a result of this oscillation some electrons may acquire sufficient energy to leave the target. Thus the existence of the photoelectric effect is in itself not surprising from a classical viewpoint. However, the principal features of the phenomenon present grave difficulties to a classical interpretation. The threshold is perhaps hardest to understand. Why should a driving force of one frequency be able to transfer enough energy to an electron to enable it to leave the target, whereas a driving force of only slightly lower frequency fails to do so, even though the amplitude of the driving electric field in the latter case might be many times greater? No plausible explanation of this enigma has ever been proposed. On the other hand, both the existence of a threshold and the qualitative form of the results are easily explained on the basis of the photon hypothesis. When an electron absorbs a single photon, it gains energy $h\nu$. If the frequency is so low that $h\nu < \phi$, then according to eq. (5) no photoelectron can leave the cathode. Hence zero current will be measured at the anode, no matter how high the accelerating voltage. If $h\nu > \phi$, photoelectrons can be ejected, and those whose kinetic energy exceeds the retarding potential will be collected at the anode. The saturation current represents the condition in which all photoelectrons are being collected.

The simplest quantitative test of Einstein's theory is based on the relation

$$\frac{1}{2} mv_{max}^2 = eV_o \quad (6)$$

in which V_o is the "extinction voltage," the potential necessary to reduce the anode current to zero. Using Eqs. (5) and (6) we can predict that a plot of V_o against frequency should be a straight line whose slope is h/e. Fig. 4 shows such a plot, taken from Millikan's paper; it is evident that a straight line fits the data quite well. The value of h obtained by Millikan from a number of runs was $(6.57 \pm .03) \times 10^{-27}$ erg seconds, in good agreement with determinations based on other experiments. (The most recent value of h is $(6.6256 \pm 0.0005) \times 10^{-27}$ erg seconds.) Moreover, the values obtained using different metals as cath-

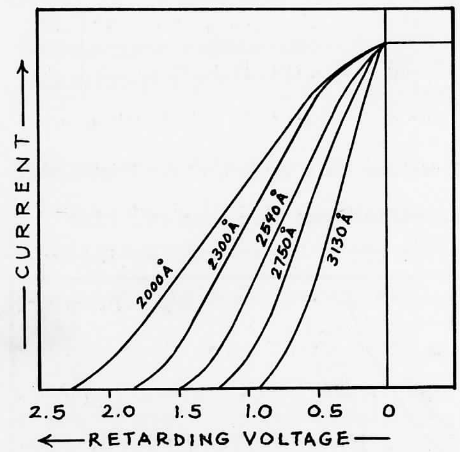


Fig. 3. Current-voltage curves for aluminum. The magnitude of the saturation current for each wavelength depends on the incident intensity. In this figure the intensities have been chosen so as to make all the saturation currents equal. (Taken from the article by O. W. Richardson and K. T. Compton, Philosophical Magazine, 24, 575 (1912).)

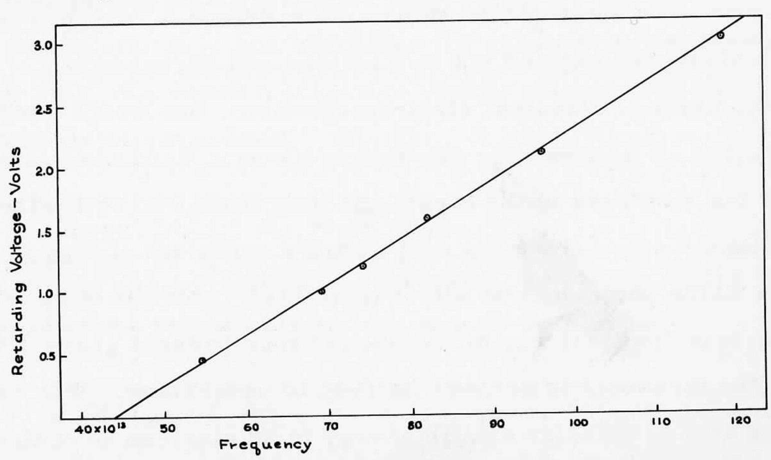


Figure 4. Determination of Planck's constant from the slope of the photoelectric cutoff voltage as a function of frequency of the incident light. Curve taken from Millikan's data for a sodium cathode.

odes were very nearly the same*.

Two additional characteristics of the photoelectric effect, which are very hard to understand from the point of view of the wave theory of light, are explained easily and simply in terms of the photon hypothesis. First, the maximum velocity of the photoelectrons for a given incident frequency is entirely independent of the intensity of the incident light. On the wave theory a more intense beam carries stronger fields, which ought to shake the electrons more vigorously and, at least occasionally, to impart greater velocities to the ones which they eject from the target. According to the photon hypothesis, on the other hand, a more intense beam merely contains a larger number of photons of the same energy, and therefore releases a larger number of photoelectrons without affecting their velocity distribution.

Further support for the photon interpretation is provided by the data on the time delay between the instant the light beam is turned on and the initial detection of a photoelectric current. According to the wave picture, the energy of the light beam is uniformly distributed in both space and time. (The sinusoidal variation is much too rapid to have any effect.) Knowing the approximate size of an atom we can estimate the rate at which energy is being deposited on the area that an atom presents to the beam. It seems reasonable to expect that the photocurrent will begin at about the time when this energy reaches a value sufficient to liberate a photoelectron. For a very weak incident beam the expected time delay can be appreciable. On the other hand, according to the photon hypothesis, the energy arrives in bunches, randomly distributed. There is some chance that the very first photon to arrive will liberate a photoelectron. (There is also some chance that no photoelectron will be liberated for a long time.) Experiment confirms the photon picture: Lawrence and Beams showed in 1928 that the photoelectric current sometimes begins in less than 3×10^{-9} seconds, even with a incident beam so weak that the expected time delay according to the wave picture is much longer.

Compton Scattering. The photoeffect is strong evidence for the existence of photons. The fact that the same value of h is obtained from experiments with different metals seems parti-

---

* The argument presented above is based on the premise that there exists a well-defined maxmum energy for the electrons in the metal. This is in fact not the case. The distribution of electron kinetic energies is studied in the field of statistical mechanics. Classical and quantum statistics predict different forms for the distribution, but according to both forms there is a definite nonzero probability that an electron's energy exceeds any specified value no matter how large. Strictly speaking, then, we should expect the curves of photocurrent (Fig. 3) to approach zero only asymptotically instead of having a well-defined intercept. Of course from an experimental point of view, once the current drops below some minimum value it cannot be distinguished from zero. And the appearance of the experimental curves indicates that the photocurrent approaches zero quite rapidly, at least at the temperature at which the experiments were carried out. Evidently the energy distribution of the electrons must be such that an "effective" maximum energy can be defined with fairly high precision. In fact, from a knowledge of the velocity distribution one can predict the precise form of the current-voltage curves; the data agree quite closely with the theoretical predictions based on quantum statistics.

cularly impressive in this regard. At the same time it is not clear what part--if any--is played by the atoms in the metal, which are intimately involved in the phenomenon. Perhaps the energy distribution of emitted electrons is a consequence of some property of the metallic atoms rather than of the quantum nature of light. Perhaps these atoms can store electrons with almost enough energy to leave the metal, so that an additional minute amount sets them off. Maybe the metal can somehow gather all the energy which falls over a wide area and transfer it to a single electron. These possibilities are admittedly somewhat farfetched, but without a complete knowledge of the interaction between matter and radiation they are difficult to refute unambiguously.

Additional and quite convincing support for the photon hypothesis is to be found in the scattering of x-rays by electrons (Compton scattering). Fig. 5 is a sketch of the experiment performed by A. H. Compton in 1923; a beam of monochromatic x-rays of known frequency strikes a target, and the frequency of the x-rays scattered at a fixed angle ϕ is measured with a crystal spectrometer (See Chapter 1 page 1-11). Fig. 6 shows a typical set of curves obtained at various scat-

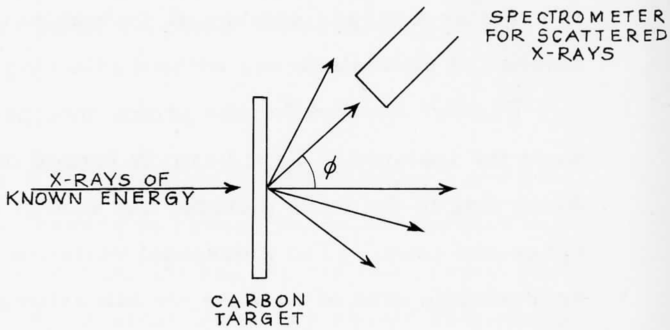


Fig. 5. Schematic diagram of Compton scattering experiment with x-rays.

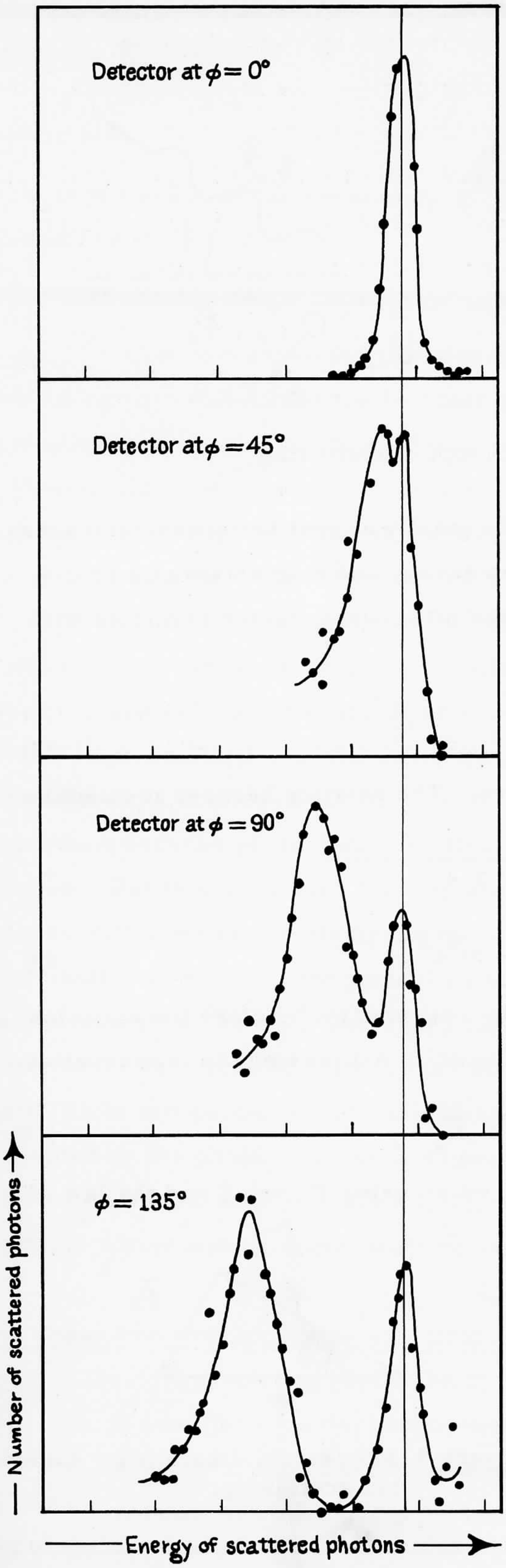


Figure 6. Results of the Compton scattering experiment with x-rays.

tering angles. The scattered x-rays are divided between two clearly separated groups--one peaked around the frequency of the incident radiation and the other peaked around a definitely lower frequency. The frequency shift increases with scattering angle, but is essentially unique at each scattering angle. It is the existence of the discrete frequency shift that leads to the photon interpretation of the effect.

Compton explained the frequency-shifted x-rays by assuming that they represent photons elastically scattered by an individual electron in the target, originally at rest. The frequency shift is then a purely kinematic result, i.e., a consequence of energy-momentum conservation. The photon is considered a completely relativistic particle whose energy and momentum are related simply by

$$E = pc \qquad (7)$$

The laws of conservation of energy and momentum imply that an incident particle always loses energy when it collides elastically with a particle initially at rest. Furthermore, the final energy of a particle scattered through a specified angle is uniquely determined by the conservation law. Thus, according to the photon hypothesis the discrete frequency shift of the scattered x-rays is explained. Moreover, the value of the frequency shift calculated from the conservation law agrees with the observations. (The necessary kinematics are worked out in Box 1.)

Box 1. Relativistic analysis of Compton Scattering*

Figure 7. Symbols for analysis of Compton scattering of a photon

A collision between an incoming photon and an electron originally at rest is represented schematically in Figure 7. The relativistic relations between energy and momentum of a photon before the collision (symbols without superscript bars) and after the collision (symbols with superscript bars) are the following:

$$E = pc$$
$$\bar{E} = \bar{p}c \qquad (8)$$

The momentum of the electron before the collision is zero. The relation between momentum and energy for the electron after the collision is

$$\bar{E}_e^2 - \bar{p}_e^2 c^2 = m^2 c^4$$

or

$$\bar{p}_e^2 = (\bar{E}_e^2/c^2) - m^2 c^2 \qquad (9)$$

Here $\bar{E}_e$ is the total energy of the electron, including the rest energy. Before the collision the electron energy is the rest energy mc^2 alone. The equation expressing the conservation of energy is

$$E + mc^2 = \bar{E} + \bar{E}_e \qquad (10)$$

The equation for the conservation of momentum may be written using Figure 8 and the law of cosines

$$\bar{p}_e^2 = p^2 + \bar{p}^2 - 2 p\bar{p} \cos \phi \qquad (11)$$

Substitute from equations 8 for p and $\bar{p}$; and from equation 9 for $\bar{p}_e^2$. Multiply the resulting equation through by c^2 to obtain

$$\bar{E}_e^2 - m^2 c^4 = E^2 + \bar{E}^2 - 2 E\bar{E} \cos \phi \qquad (12)$$

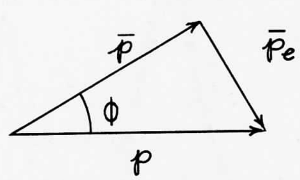


Figure 8. Momentum diagram for Compton scattering.

The energy $\bar{E}_e$ of the electron after the collision is of no interest to us. Solve equation 10 for $\bar{E}_e$ and substitute this value into equation 12. Expand and simplify to obtain

---

* See Physics, a New Introductory Course, 1966 edition, Part III, Chapter 6, page 20; or 1965 edition Part III, Chapter 21, page 18.

$$Emc^2 = \bar{E}(E + mc^2 - E\cos\phi)$$

Solve for $\bar{E}$:

$$\bar{E} = \frac{E}{1 + (E/mc^2)(1 - \cos\phi)} \qquad (13)$$

This is the equation for the energy $\bar{E}$ of a photon emerging at an angle ϕ, when the incident photon had energy E.

Compton explained the presence of the unshifted peak in the scattered radiation by supposing that sometimes the incident photon interacts with the atom as a whole rather than with a single electron which acts as though it were free. In an elastic collision with an atom at rest a photon still loses energy. However, because the mass of the atom is many times greater than the mass of an electron, the energy loss (and consequently the frequency shift) are so small as to be undetectable. (Equation (13) of Box 1 with m now thought of as the mass of entire atom.)

It should be mentioned that the mere detection of frequency-shifted x-rays is not an argument for the photon nature of light. Classical electromagnetic theory also predicts a frequency shift. If one solves the problem of a free electron acted on by an oscillating electric field, the result is an oscillation of the same frequency superposed on a uniformly accelerated motion. Therefore the scattered radiation which one observes in a particular direction has a different frequency because of the Doppler effect which occurs when source and observer are in relative motion. But in such a case the Doppler shift does not have a fixed value; in fact it should increase with time as the electrons continue to accelerate. (See the problems.) Therefore the distribution of energy (frequency) of scattered light should form a single wide peak--according to this classical theory. The sharp displaced peak of Fig. 6 is in marked contrast to this classically predicted distribution. Thus classical theory, even though it does predict a frequency shift, is in strong disagreement with the observed nature of the Compton effect. The agreement provided by the photon picture is impressive. Notice that only kinematics is used to explain the effect. No assumption is needed concerning the nature of the photon other than that it is a particle whose energy-momentum relationship is given by Eq. (7).

2. Other properties of photons: linear momentum and angular momentum.

Classical electromagnetic theory predicts that

1. A beam of light carries energy.
2. A beam of light carries linear momentum.
3. A beam of circularly polarized light (see Chapter 3) carries angular momentum.

All these are statements of classical physics. There is no mention of photons, of Planck's constant, or of quantum states. The three properties are derivable from Maxwell's equations, developed in all essentials a quarter of a century before the first steps of quantum physics were taken.*

* For the development of the three statements from Maxwell's equations, see for instance, J. D. Jackson Classical Electrodynamics, John Wiley and Sons, Inc. 1962, pages 189, 191, 201.

All three predictions stated above are confirmed by experiment. Anyone standing in direct sunlight does not need to be told that light carries energy. The linear momentum carried by a beam of light can be observed by letting the beam be absorbed by one vane of a sensitive torsion balance (Fig. 9). The angular momentum carried by a beam of circularly polarized light was measured in 1932 by R. A. Beth.* More recently a similar experiment has been carried out with circularly polarized microwaves--electromagnetic radiation of frequency much lower than that of light.** This experiment is diagrammed in Fig. 10. A circularly polarized beam is

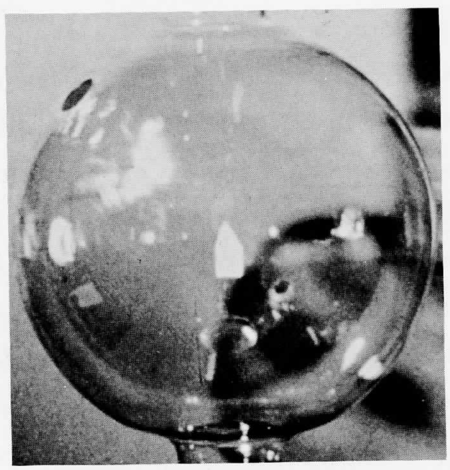


Figure 9. Mirror suspended from a thin fiber used as a torsion balance to detect the linear momentum carried by an incident beam of light. Photograph from the Physical Science Study Committee film "The Pressure of Light"

absorbed by a cylinder which is the rotating member of a torsion pendulum. The angular momentum of the beam absorbed by the cylinder causes the cylinder to rotate. This rotation is opposed by the restoring force of the suspension. In equilibrium a static deflection of ten degrees is easily obtained.

Quantitative results obtained in experiments such as these agree with theoretical predictions derived from Maxwell's equations, namely,

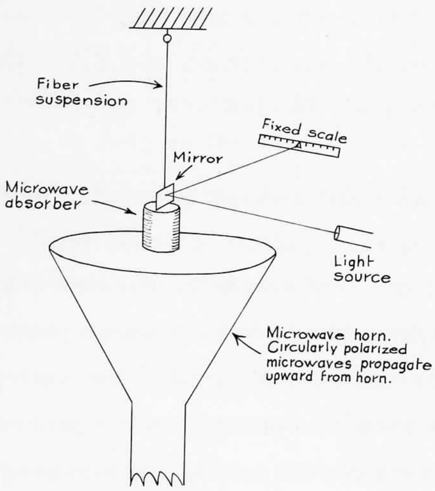
1. Let a beam of electromagnetic radiation carry energy W per square centimeter per second into an absorber.
2. Then linear momentum W/c is added to the absorber per square centimeter per second.

---

* R. A. Beth, Physical Review, 50, 115 (1936).

** Richard B. Anderson and Joseph S. Ladish, Senior Thesis, M. I. T. June, 1965 (unpublished). A filmed demonstration is being produced by Educational Services, Inc., Watertown, Mass. See also P. J. Allen, American Journal of Physics, 34, 1185 (1966).

3. If the incident beam is circularly polarized, then the angular momentum (along the direction of the beam) carried into every square centimeter of the absorber every second by the beam is $\pm W/2\pi\nu$, where ν is the classical frequency of the radiation.



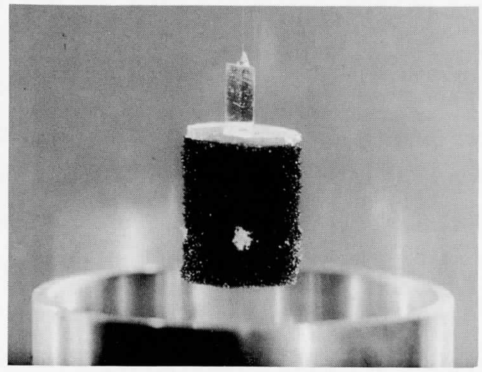


Figure 10. Experiment to detect angular momentum of circularly polarized microwaves. Diagram from M.I.T. Senior Thesis (1965 unpublished) by Richard B. Anderson and Joseph S. Ladish. Photograph is from E.S.I. film "The Angular Momentum of Circularly polarized Radiation."

So much for the well-defined classical predictions regarding linear and angular momentum of light. How are we to interpret these results in terms of the alternative point of view, that light is composed of photons? One is tempted to say (hypothesis 1) that each photon of frequency ν carries linear momentum, $h\nu/c$, and each photon in a circularly polarized beam carries angular momentum $\pm h/2\pi\nu$ along the beam direction. These assertions are indeed correct, but they surely are not proved by the experiments referred to above. All that is determined by such classical experiments is the average linear and angular momenta carried by the photons. One can devise other hypotheses consistent with the same classical results, for example: (hypothesis 2) linear (angular) momentum is a property resulting from the cooperative behavior of photons, not a property of individual photons; (hypothesis 3) some photons carry linear (angular) momentum while others do not; (hypothesis 4) some photons in a monoenergetic beam carry more linear (angular) momentum than other photons. Hypotheses (1) through (4) are statements about photons. Therefore no classical theory or experimental result can help us decide among these or other alternatives. We must look for help to experiments in which the quantum nature of photons is an essential feature.

One way to test for possible cooperative phenomena between photons (hypothesis 2) is to reduce the intensity of light in the beam. With a smaller flux of photons the opportunity for cooperation between photons is reduced. Nevertheless, although the energy flux W is smaller for a less intense beam, the average linear momentum is still observed to be W/c and the average angular momentum (for a circularly polarized beam) is still $\pm W/2\pi\nu$ independent of beam intensity. Is it possible (hypotheses 3 and 4) that different photons of the same energy carry different amounts of linear momentum? The analysis of the Compton experiment in the preceding section argues strongly against such a possibility. For if some of the incident photons of energy $h\nu$ carried a momentum other than $h\nu/c$, their frequency shift after scattering from an electron would be different from the result (Eq. 13) calculated on the assumption $p = E/c$. One would then predict the presence of more than one group of frequency-shifted photons, contrary to observation.

The fact that each circularly polarized photon carries an equal amount of angular momentum in the direction of motion is not so easy to demonstrate as is the equality of linear momentum. Assuming that the law of conservation of angular momentum can be carried over from classical mechanics, we can obtain some evidence by analyzing phenomena in which photons interact with systems whose angular momenta are known or can be measured. One such phenomenon is the emission of radiation by an atom. If a single photon is emitted in the process, its angular momentum is, presumably, the vector difference between the initial and final angular momenta of the emitting atom.* It is found that transitions in which the initial and final atomic angular momenta are both zero never occur. This strict selection rule strongly suggests that photons of zero angular momentum do not exist; it does not prove that they do not exist, nor does it prove any equality among angular momenta of photons. However, more detailed study indicates that when an atom emits a photon in, say, the z direction, the z component of the atom's angular momentum always changes by $\pm h/2\pi$. This is a strong indication that the component of the photon's angular momentum along its direction of motion is indeed $\pm h/2\pi$.

In summary, then, there is good evidence that a beam of electromagnetic radiation of classical frequency ν behaves as a stream of photons, each with energy $h\nu$, linear momentum $h\nu/c$, and (if circularly polarized) angular momentum $\pm h/2\pi$ along the direction of motion.

The present chapter has dealt with the existence of photons and with the properties which the particle model of light assigns to photons. The next few chapters consider specifically the polarization of photons as an example of the states of a quantum mechanical system. Many of the results fundamental to the formulation of quantum mechanics will in this way be laid bare in terms of tabletop experiments. As we have already emphasized, the central feature of this treatment is neither the photons nor the particular experiments themselves, but rather the concepts which arise in the analysis. These concepts are useful in understanding a wide range of atomic and subatomic phenomena.

---

* This statement will be qualified after more is known about the quantum properties of angular momentum.

CHAPTER 3. PHOTON STATES*

1. Introduction

In Chapters 3 through 5 we analyze some experiments with polarized light in terms of the photon concept. Classical wave optics accounts easily and simply for the results of these experiments as long as the light beams are sufficiently intense.** A similar accounting using photons forces one into a point of view which is radically different from that of classical optics and, to begin with, somewhat more complicated. Why bother with the more complicated photon interpretation? For two reasons: one scientific and one pedagogic. In the first place other experiments--such as those already discussed in Chapter 2--require the photon concept for their interpretation. Therefore one is interested in finding out whether all experiments with light can be interpreted consistently in terms of photons. In addition, the kind of analysis used here for photons can be applied with very little modification to quantum experiments with other particles, such as protons, electrons, atoms, and molecules. The results of many experiments with these other particles have no classical explanation analogous to the wave theory explanation of polarization phenomena. The study of polarized light makes a convenient bridge between classical ideas and the concepts of quantum mechanics.

2. The classical description of linear polarization

A useful idealization in classical electromagnetic theory is the "plane wave" of radiation--a particular solution of the wave equation in which the "wave fronts" (surfaces of constant phase) are planes perpendicular to the direction of propagation. Everywhere on each such plane the electric field $\vec{E}$ has, at a given instant of time, the same magnitude and direction. The magnetic field $\vec{B}$, as in any wave solution in free space, is everywhere perpendicular both to $\vec{E}$ and to the propagation direction. If $\vec{E}$ is measured in e.s.u. (statvolts per centimeter) and $\vec{B}$ is measured in gauss, the magnitudes of the two fields in a plane wave are everywhere numerically equal.

An ideal plane wave has infinite width and extends to infinity in both directions along the line of propagation. No real (i.e., finite) source can generate such a wave. All real beams have finite cross sections. In addition, most beams of light are not coherent over large distances: surfaces of given phase are not planes and are not equally spaced to infinity in both directions. Nevertheless the concept of a plane wave is a useful one for analyzing many experiments.

---

* The authors acknowledge the influence on this chapter of a movie script written by Prof. A. P. French.

** An excellent brief introduction to the classical theory and practical applications of polarized light is presented in the book Polarized Light by William A. Shurcliff and Stanley S. Ballard, Momentum Book #7, D. Van Nostrand Co., Princeton, 1964.

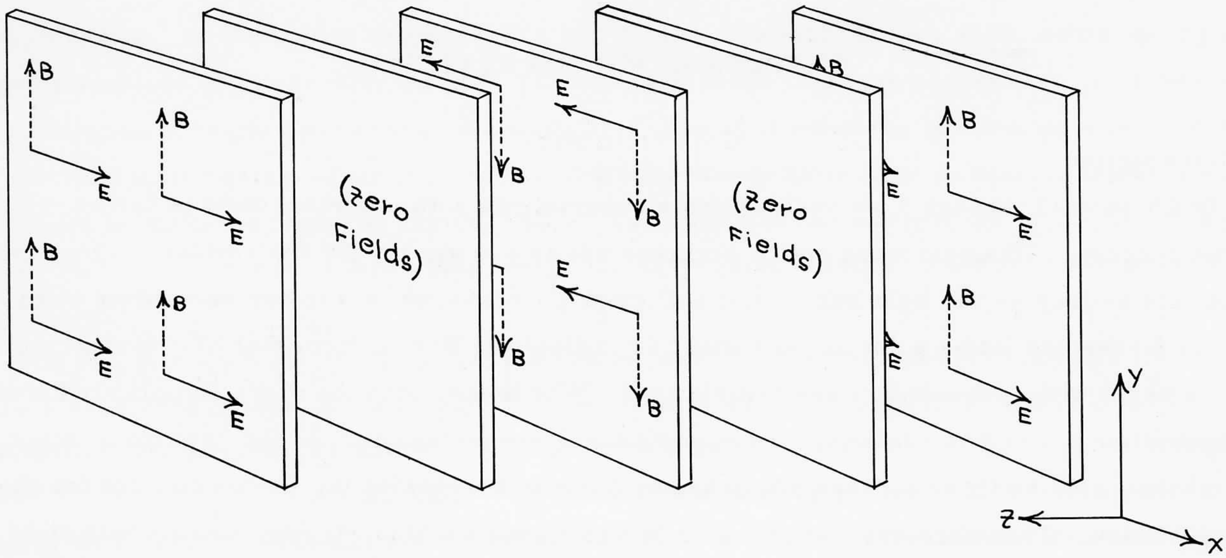


Figure 1. Portion of a linearly polarized plane wave at a fixed time, showing electric fields (solid arrows) and magnetic fields (dashed arrows) in selected planes of constant phase. The entire pattern propagates in the positive z direction.

A linearly polarized plane wave is one in which the electric vector is everywhere parallel to a line fixed in space. The magnetic vector points in a direction perpendicular to this fixed line. Figure 1 shows at a given instant of time a linearly polarized beam propagating in the positive z direction. The analytical expression for this wave is

$$\begin{aligned} E_x &= B_y = A\cos(kz - \omega t) \\ E_y &= E_z = B_x = B_z = 0 \end{aligned} \qquad (1)$$

where the angular frequency ω and the wave number k are related to the frequency ν and the wave length λ:

$$\begin{aligned} \omega &= 2\pi\nu \\ k &= 2\pi/\lambda = \omega/c \end{aligned} \qquad (2)$$

Call the line parallel to the electric vector the polarization axis. (It is conventional to define the polarization axis in terms of the electric vector rather than in terms of the magnetic vector or some combination of the electric and magnetic vectors.) The polarization axis of the wave in Figure 1 is the x-axis; this wave is said to be x-polarized. More generally, the polarization axis of a linearly polarized wave that propagates in the z direction can be any line in the xy plane.

A simple practical method for producing light linearly polarized along a prescribed axis is based on the fact that waves of different polarization have different reflection characteristics at the interface between two dielectrics. Figures 2a and 2b show a beam of light incident from air onto a flat sheet of glass. The direction of the beam and the normal to the surface determine a

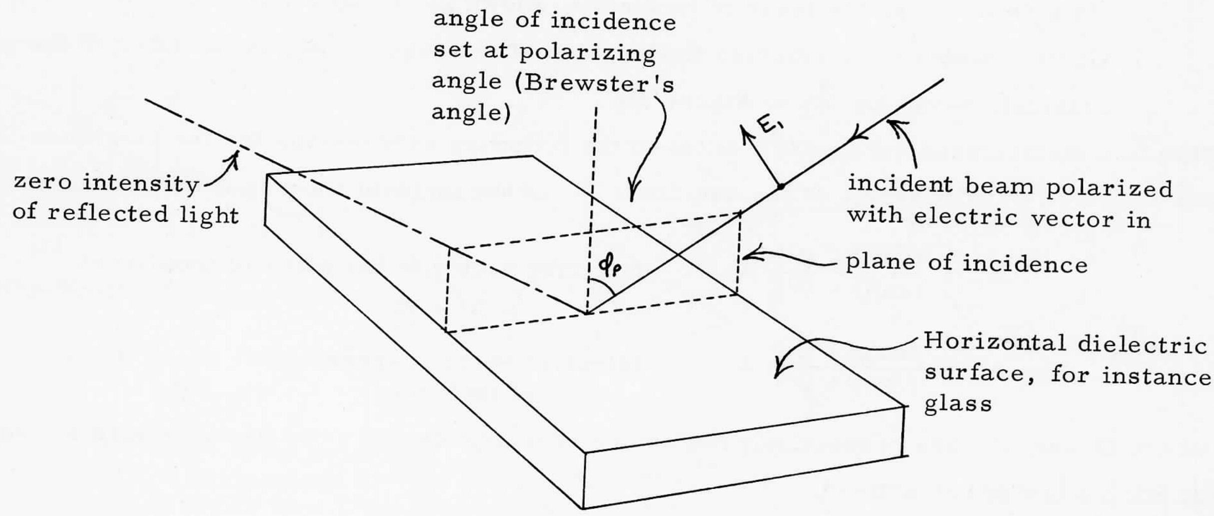


Figure 2a. When light linearly polarized in the plane of incidence strikes glass at Brewster's angle of incidence, no light is reflected. There is of course a refracted beam (not shown in the figure).

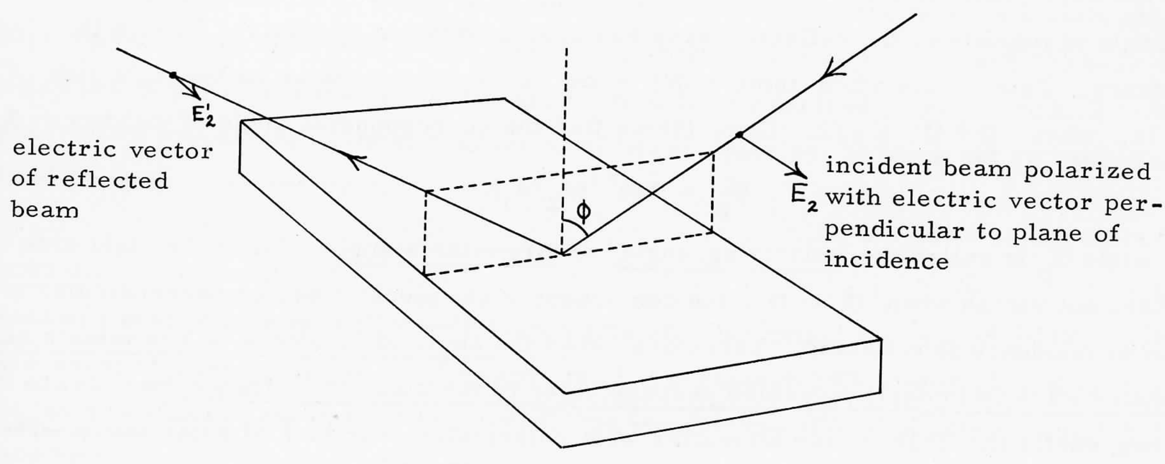


Figure 2b. When light linearly polarized perpendicular to the plane of incidence strikes glass at any angle of incidence, some light is reflected.

plane called the plane of incidence. We consider two possible directions for the electric vector of the incident wave:

i) electric vector in the plane of incidence, shown as $\vec{E}_1$ in Figure 2a;

ii) electric vector perpendicular to the plane of incidence (i. e., in the plane of the glass surface), shown as $\vec{E}_2$ in Figure 2b.

Standard electromagnetic theory* leads to the following expressions for the amplitude A' of the reflected wave in terms of the amplitude A of the incident wave in these two cases:

i)
$$A' = \frac{\tan(\phi - \phi')}{\tan(\phi + \phi')} A \qquad \text{(electric vector in the plane of incidence)} \tag{3}$$

ii)
$$A' = \frac{\sin(\phi - \phi')}{\sin(\phi + \phi')} A \qquad \text{(electric vector perpendicular to the plane of incidence)} \tag{4}$$

where ϕ and ϕ' are respectively the angles of incidence and refraction, related to one another by Snell's law of refraction

$$n_1 \sin \phi = n_2 \sin \phi' \tag{5}$$

Here n_1 and n_2 are the indices of refraction of the air and glass respectively. (For normal incidence one must use (5) and take the limit of (3) and (4) as ϕ and ϕ' approach zero.)

The equations above show us how to produce linearly polarized light (Figure 3). The The electric vector of any plane wave incident on the glass can be analyzed into components parallel and perpendicular to the plane of incidence. Equations (3) and (4) describe the fate of each of these components upon reflection. The resultant electric field of the reflected wave is then the superposition of the reflected components. Equation (3) tells us that, for one particular angle of incidence, the reflected wave has zero component of electric field in the plane of incidence. This occurs when $\tan(\phi + \phi') = \infty$ in the denominator on the right side of (3), that is, when $\phi + \phi' = \pi/2$. Using (5) we find the corresponding angle of incidence ϕ_p to be

$$\phi_p = \tan^{-1}(n_2/n_1) \tag{6}$$

The angle ϕ_p is called the polarizing angle or Brewster's angle. Since the right side of eq. (4) does not vanish when $\phi = \phi_p$, the component of the electric vector perpendicular to the plane of incidence is reflected. Therefore when the angle of incidence is Brewster's angle the reflected wave is linearly polarized perpendicular to the plane of incidence (parallel to the reflecting surface). This is true no matter what polarization--or lack of polarization--the incident wave has (Figure 3). (In the special case when the incident wave is polarized parallel to the plane of incidence no light will be reflected: the incident beam will be entirely refracted into the glass.)

Polarization by reflection yields a plane wave whose axis of polarization is known. Even when the angle of incidence differs from Brewster's angle, the beam reflected from a dielectric surface is partially polarized, since the component of the electric field perpendicular to the

* Optics by Bruno Rossi, Addison-Wesley Publishing Company, 1957, p. 366 ff.

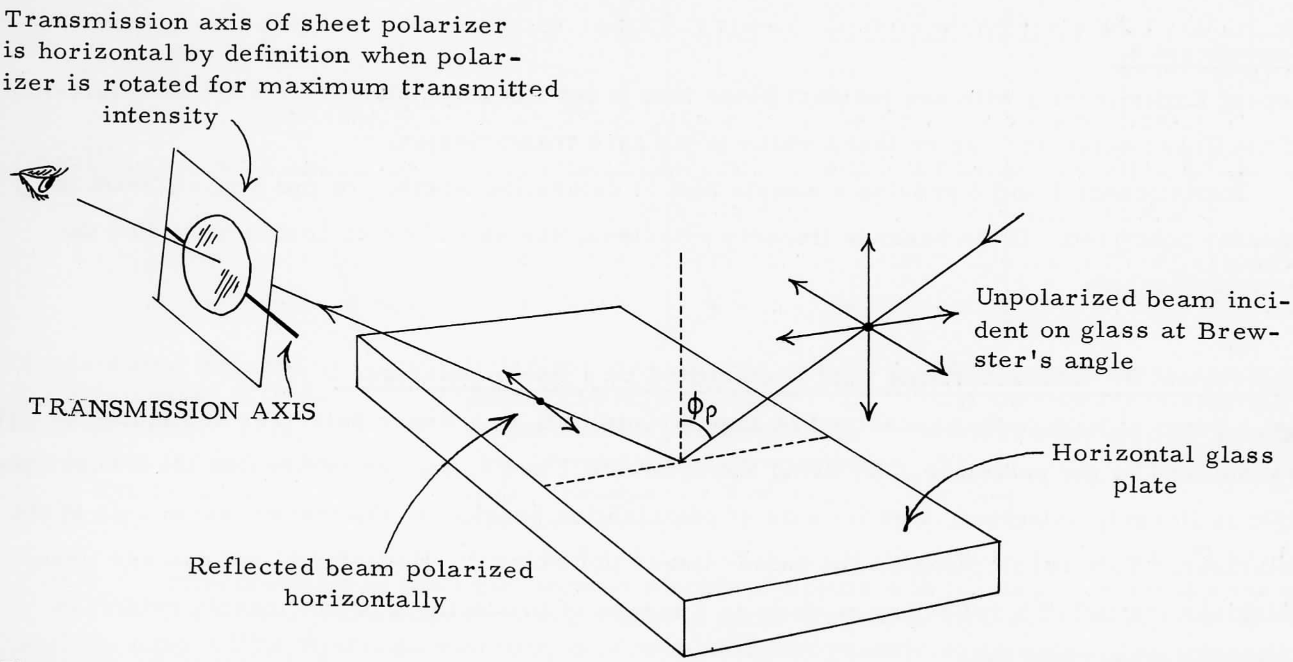


Figure 3. Polarization by reflection (Section 2). Determination of transmission axis of linear polarizer (Experiment 1 of Section 3).

plane of incidence is reflected more strongly than the component parallel to the plane of incidence.

3. Experiments with linearly polarized light.

The first three experiments to be described in this section establish the properties of a material called linear polarizer, which is commercially available in the form of plastic sheet. These experiments show that the material can be used both to test a given beam for linear polarization and to produce a linearly polarized beam with any desired polarization axis. A fourth experiment makes use of these properties of linear polarizer to bring out an essential property of linearly polarized light.

Experiment 1.

Let a beam of linearly polarized light with known axis of polarization impinge at normal incidence on a sheet of linear polarizer. The beam may, for example, have been polarized by reflection as described in the preceding section. On rotating the linear polarizer in its own plane, one finds that the intensity of the transmitted beam varies. For a certain orientation the transmitted intensity is, ideally, zero (in practice, almost zero). When the polarizer is rotated by 90 degrees from the orientation that gives zero transmission (Figure 3), the intensity of the transmitted beam is found to be a maximum. With the polarizing sheet oriented for maximum transmission, we can scratch on it a line parallel to the known polarization axis of the incident light and call this line the transmission axis of the linear polarizer. We suppose hereafter that the transmission axis has been so indicated on all linear polarizers used in our experiments.

Experiment 2.

Repeat Experiment 1 with any incident beam that is not linearly polarized. Then no orientation of the linear polarizer can be found which gives zero transmission.

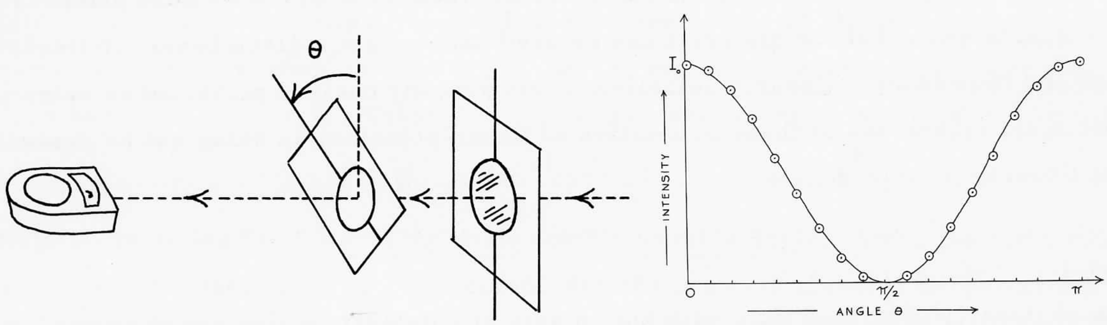
Experiments 1 and 2 provide a simple test to determine whether or not a given beam is linearly polarized. If the beam is linearly polarized, the experiments enable us to find the polarization axis.

Experiment 3. Characterizing light transmitted by a linear polarizer

Let a beam of light (either polarized or unpolarized) fall on a linear polarizer and study the light transmitted by the polarizer. By using Experiments 1 and 2, we can verify that the transmitted light is linearly polarized, with its axis of polarization parallel to the transmission axis of the polarizer. This result justifies the name "linear polarizer." Henceforth we can use linear polarizer instead of a reflecting surface as a means of producing a beam linearly polarized along any prescribed axis.

Experiment 4.

Let a linearly polarized beam produced by polarizer A (Figure 4) impinge on polarizer B whose transmission axis makes an angle θ with that of A. When θ is varied, by rotating either polarizer in its own plane, the intensity of the beam emerging from B varies as $\cos^2\theta$. Figure 4 diagrams the apparatus and displays the results for this fundamental experiment.



Intensity-measuring device, for instance photographic exposure meter.

Sheet polarizer B

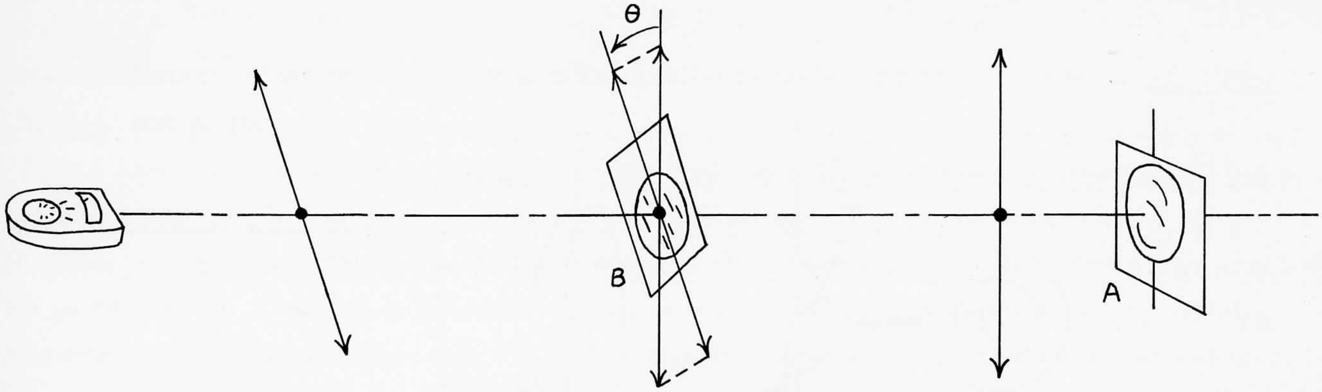
Sheet polarizer A

Results of an experiment like the one pictured at left for several values of the angle θ. Circled dots are experimental points. Solid curve is

$$I = I_0 \cos^2\theta$$

predicted by classical optics-- see Figure 5.

Figure 4. Transmission of linearly polarized light by a sheet polarizer (Experiment 4 of Section 3).



Transmitted intensity is fraction $\cos^2\theta$ of incident intensity. This prediction is plotted as solid line on graph of Figure 4.

Polarizer B transmits component of electric vector projected along transmission axis (magnitude of component is the fraction $\cos\theta$ of incident amplitude).

Maximum positive and negative component of electric field transmitted by polarizer A.

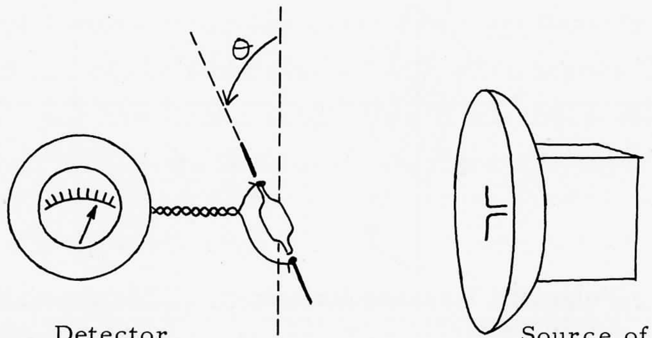
Figure 5. Interpretation of the experimental results of Figure 4 in terms of classical wave optics. For the sake of simplicity the assumption is made that the polarizer is ideal, that is, it absorbs none of the light with vibration direction along the transmission axis. The intensity of light is proportional to the square of magnitude of the electric vector.

Classical interpretation of experiments. Experiments 1 to 4 may be interpreted on the basis of classical wave optics if we suppose that the structure of the linear polarizer is such that it transmits only the electric field component parallel to its transmission axis (Figure 5). When a wave with an electric field $\vec{E}$ along some other axis is incident on the polarizer, only the component of $\vec{E}$ along the transmission axis is transmitted. The component of $\vec{E}$ perpendicular to the transmission axis is absorbed. We are not concerned here with the detailed structure of the polarizer that leads to this property. (See the caption of Figure 7 for a brief statement on this structure.) The results of Experiments 1 to 3 follow directly from the property assumed for linear polarizer. In Experiment 4 the electric field of the transmitted wave varies as $\cos\theta$ (Figure 5). According to electromagnetic theory the intensity (energy flow per unit area per unit time) is proportional to the square of the electric field. Therefore the intensity of the emergent beam is proportional to $\cos^2\theta$.

Linear polarization is exhibited not only by visible light, but also by radiation throughout the electromagnetic spectrum. For example, devices can be constructed whose effect on microwaves are analogous to the effects of linear polarizing sheet on visible light. Figures 6 and 7 summarize experiments with microwaves similar to the experiments with polarized light described above.

4. Linear polarization as a quantum state

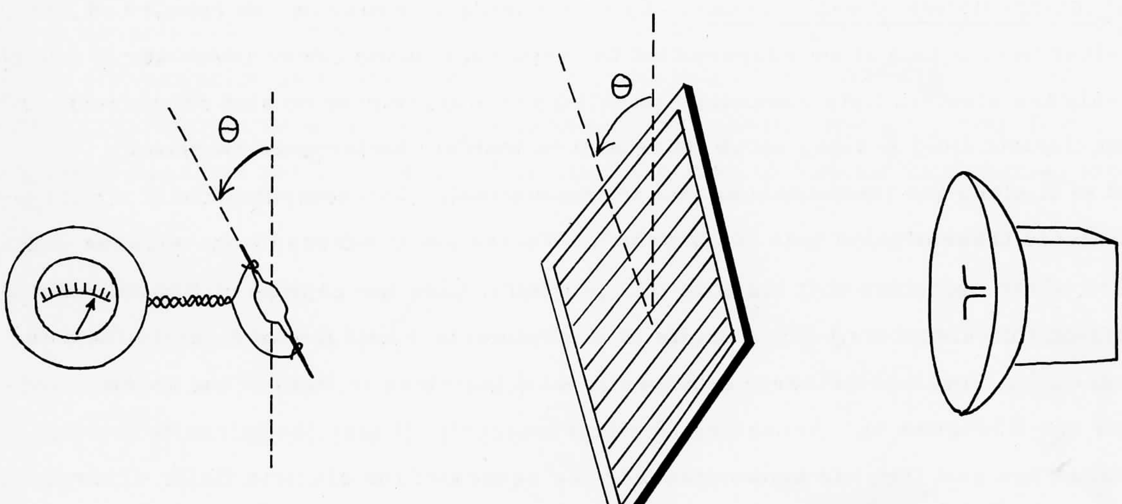
We now re-examine the experiments described in the preceding section, considering the light beam as a stream of photons rather than as a classical wave. It is reasonable to call the photons in a linearly polarized beam linearly polarized photons. Photons that emerge from a



Detector
(solid state diode and
dc galvanometer)

Source of
polarized microwaves

Figure 6. Direct measurement of polarization direction of electromagnetic waves at microwave frequency. Detector reading is maximum when detector wire is parallel to the source antenna (when angle $\theta = 0$); zero when detector wire is perpendicular to the source antenna (when angle $\theta = 90^{o}$).



Transmitted wave polarized perpendicular to wires in filter. Amplitude is fraction $\cos\theta$ of amplitude of beam incident on filter.

Microwave "sheet polarizer" (parallel wires stretched across insulating frame).

Source of polarized microwaves

Figure 7. An experiment with a "microwave sheet polarizer." The polarizer transmits the component of the electric field perpendicular to the wires. The component parallel to the wires induces currents in the wires which cancels this parallel component. Optical sheet polarizer works much the same way: iodine dye on parallel-oriented long-chain molecules allow conduction along these molecules but not perpendicular to the molecules.

linear polarizer whose transmission axis is parallel to the x-axis will be called x-polarized photons, and so on.

We assert that linear polarization (together with energy and direction of motion) defines a quantum state for photons, in the sense discussed in Chapter 1. Recall the principal characteristics of a quantum state. First, there must exist an unambiguous test to determine whether the particles in a given set do or do not belong to the class that defines the state in question. Experiments 1 and 2 above provide the required test: the photons in a given beam are in the quantum state of x-polarization if and only if the beam is transmitted with undiminished intensity by an ideal linear polarizer with its transmission axis parallel to the x-axis. In this definition an ideal linear polarizer is specified because commercial polarizing sheet transmits only 80 to 90 percent of incident light linearly polarized along its transmission axis. An ideal linear polarizer is defined as one that transmits 100 percent of the light polarized along its transmission axis and none of the light polarized perpendicular to this axis. In Section 6 we shall introduce a calcite polarizer that can be made more nearly ideal.

The second characteristic required of a quantum state is that there be no criterion that serves to subdivide the class any further. Now x-polarized photons can certainly have different energies and different directions of motion. In the present discussion we are concerned almost exclusively with polarization aspects of photon states and shall generally not trouble to specify the energy or direction of motion. However, it is important to keep in mind that information concerning these characteristics is part of the full definition of photon quantum state.

The class of x-polarized photons, with energy E, travelling in the z-direction, cannot (as far as we know) be further subdivided. This class therefore constitutes a strict quantum state. This is not an obvious fact, but is a conclusion reached only after much experimentation and is subject to revision if additional evidence demands it. In an analogous manner one can define the state of y-polarization, or of linear polarization along any axis perpendicular to the direction of motion of the photons.

Notice that we have not defined an x-polarized photon as one whose electric field points along the x-axis. No experiment reported in Chapter 2 enables us to measure the electric field of a single photon. Detection of a photon in those experiments is signaled for instance, by a pulse of current from a photomultiplier tube or the exposure of a single silver bromide crystal in a film emulsion. A more complete theory of photons, in fact, shows that the electric field is a classical concept that can be realized only when a great number of photons are present. However, the limited meaning of "electric field" does not limit in anyway the experiments we use to identify photon states. With a linear polarizer we distinguish experimentally between a class of photons that we choose to call x-polarized, and a different class that we choose to call y-polarized. That is all that is required for the description of photon polarization states. This description works just as well for a weak beam, for which an electric field cannot be defined, as for an intense beam for which a classical wave description may be valid.

Now apply the photon concept to interpret experiment 4 above (Figure 4). In that experiment the transmission axes of polarizers A and B make an angle θ. The intensity of the light that emerges from polarizer B is a fraction $\cos^2\theta$ of that which enters B. The classical explanation is that polarizer B transmits only the component of electric field along its transmission axis. For a photon explanation one might suppose--incorrectly--that linear polarizer B somehow splits each photon incident on it into two, transmitting only the member of each pair that is "polarized along the transmission axis." This explanation is defeated by the incorrectness of its predictions. If "part of each incident photon" is transmitted, then there are the same number of photons in the transmitted beam as in the incident beam. In order to explain the measured lower intensity of the emergent beam, one is forced to say that each photon emerging from B carries (on the average) less energy than does a photon incident on B. Therefore, according to the basic relation $E = h\nu$, the frequency of the emergent beam must be correspondingly smaller than that of the incident beam. This prediction is incorrect. By placing a spectrometer first in the beam incident on polarizer B and then in the transmitted beam, one verifies that no matter how the polarizers are oriented the frequencies of the two beams are exactly the same. It follows that each photon in the emergent beam has exactly the same energy as each photon in the incident beam. The observed decrease in intensity means, then, that the emergent beam carries fewer photons per second than does the incident beam. We conclude that a fraction $\cos^2\theta$ of the photons incident on polarizer B are transmitted by it, while a fraction $\sin^2\theta$ of the incident photons are absorbed. We say that, for photons in the state determined by the transmission axis of polarizer A, the probability of transmission by an ideal polarizer B is $\cos^2\theta$, where θ is the angle between the transmission axes of polarizers A and B.

Notice that photons transmitted by polarizer B are not in the same quantum state as are the photons that enter B. The photons that enter B are in the state defined by the transmission axis of polarizer A, as can be verified by passing them through a linear polarizer whose axis is parallel to that of A. The photons emerging from B are similarly verified to be in the quantum state defined by the transmission axis of B. This change of state upon measurement is a very important general feature of quantum mechanics. Suppose we are given a set of particles (say photons) and test them to determine whether or not they are in some specified quantum state (say the state of x-polarization). If the result of the test is positive, the particles after the test are still in the same quantum state. An additional test will give the identical result. But suppose the test is negative, i.e., suppose we learn that the original particles were not in the state for which we tested them. Then the test itself has necessarily changed the state of the particles. In our example we test for x-polarization a beam of photons that are not x-polarized. Some of the photons are absorbed and we are left with a beam of lower intensity which consists of photons that are x-polarized. In systems amenable to classical description we can determine the configuration while disturbing the system by an arbitrarily small amount.

In systems describable only by quantum mechanics we always change the state of the particles whenever we carry out an experiment on them, except when we test them for the quantum state in which they happen to be.

Before going on to discuss probability we make a cautionary remark about the identifiability of individual photons. In the often-used phrase "the photons transmitted by polarizer B," the word "transmitted" seems to imply that the photons emerging from the polarizer are in some sense "the same photons" that entered the polarizer. Yet there is no way to verify this sameness. In Experiment 4 all the photons entering polarizer B are in the same state; similarly those photons emerging from B are in the same state (though perhaps this state is different from that of the incident photons). We have no way of knowing whether or not the photons that emerge from B are the same ones that enter. When this question is pursued further by analyzing the details of interaction between incident photons and the atoms of the polarizer, the picture implied by quantum electrodynamics involves numerous absorptions and re-emissions of photons as the beam proceeds through the material. (The same kind of absorption and re-emission can be used to account for--among other things--the slower speed of light in glass than in a vacuum.) Insofar as such a picture can be given an everyday meaning, we would have to conclude that the emerging photons are "different" from those that enter the material. We cannot concern ourselves here with the details of these interactions. In fact, such knowledge is not necessary to our present purpose. All we measure in polarization experiments is the relative intensities of various beams.

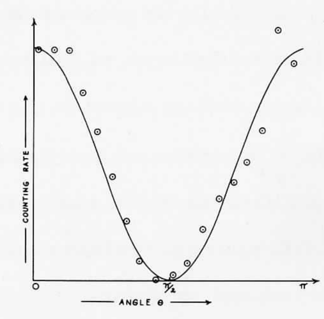
5. Probability in quantum physics

Thus far we have dealt mainly with statements that apply collectively to all the photons in a beam. To interpret Experiment 4, we said that a fraction $\sin^2\theta$ of the photons incident on polarizer B is absorbed, and a fraction $\cos^2\theta$ is "transmitted" in the quantum state defined by the transmission axis of the polarizer. Can we say anything about what happens to individual photons? Statements about individual photons are statistical in nature: one says there is a probability $\sin^2\theta$ for any given incident photon to be absorbed, and a probability $\cos^2\theta$ for it to be transmitted in the new polarization state. Beyond stating such probabilities, we cannot predict what will happen to any particular photon; it seems to be a matter of chance. This statement is an extremely important one; it denotes a fundamental difference between the quantum and classical descriptions of nature. The classical description is completely deterministic: given sufficient information about a classical system, one should be able, at least in principle, to predict the exact outcome of any experiment on that system. We are now saying that the specification of a quantum state constitutes the maximal information one can have about a collection of identically prepared systems, and only statistical predictions can be made about most experiments carried out on them.

Probability and the statistical nature of theory and experiment are central features of quantum physics. Historically this conclusion did not meet with immediate and universal acceptance; on the contrary it has engendered much debate among scientists and among philosophers, debate that continues to the present day. Many have been dissatisfied with a probabilistic description of the universe. Some argued as follows: "Perhaps if we really knew everything about the system, that is if we could find additional properties of the system ("hidden variables") not specified by the quantum state, then we could predict with certainty the outcome of each individual event." Einstein, for one, was strongly opposed to the probabilistic interpretation. "God," he said, "does not play dice with the universe." ("Der Herr Gott würfelt nicht.") But after more than forty years, no one has been able to display any "hidden variables" that make possible a classical type of deterministic interpretation. The present consensus among scientists is strongly in favor of the probabilistic interpretation of atomic and sub-atomic events. But the probabilities for any experiment with many events are very well determined, and it is the business of quantum physics to measure and calculate these probabilities.

That experiment 4 (Section 4, Figure 4) has statistical aspects can be demonstrated directly. Suppose that the experiment is carried out with an extremely weak incident beam, one so weak that only a few photons, on the average, are present in the region of the apparatus. It is necessary to use a photomultiplier tube as a detector of sufficient sensitivity to monitor the low-intensity beam. The photomultiplier records a count only occasionally; the counts occur in an irregular pattern, with a typical statistical distribution in time. As the total number of counts becomes large, the average number of counts per unit time approaches the fraction $\cos^2\theta$ of the average counting rate observed when the polarizer is absent. But during short time interval the counting rate is subject to the fluctuations characteristic of statistically independent events.* Figure 8 shows the result of an experiment of this sort, with a weak beam and measurements confined to short time intervals. The $\cos^2\theta$ dependence is seen to be only approximately fulfilled. With light of low intensity the statistical average can be more accurately verified by recording large numbers of events in an experiment of long duration. In practice one generally uses beams of high intensity, because photons do not interact appreciably with one another--a feature of their behavior that recommended photons to us in the first place. Therefore the intensity of the beam is immaterial as long as the total number of counts is the same. With high intensity the $\cos^2\theta$ dependence can be accurately verified in a short time.

Figure 8. Results of experiment similar to that of Figure 4, but with a very weak beam and a photomultiplier tube detector. Each point is based on 100 counts recorded by the detector. Statistical fluctuations are evident.



* See, for example, Physics, A New Introductory Course, Vol. 1, Chapter 9.

The probabilistic interpretation is very much a part of our definition of a quantum state. Notice that even though the polarization state is a state of the individual photon, there is no way to determine the state of a single photon. For if we let the photon impinge on, say, an x-polarizer and observe that a photomultiplier on the other side registers a count, we cannot conclude that the original photon was x-polarized. It may have come from a beam of photons linearly polarized along any axis (except the y axis), or from a beam of circularly polarized photons (see Section 7 below), or from any one of a number of other beams. For a meaningful state determination we must verify that the beam emerging from the polarizer is just as intense as the beam that enters. The measurement of intensity necessarily involves the detection of many photons. In general one determines the quantum state of each member of a large set of identically-prepared particles. Such a determination involves making the same measurement on many members of the set. This is why the identity of atomic particles plays such an important role in quantum physics. The measurements are made on individual particles one at a time. With an intense beam of photons we do not need a photomultiplier tube, and we may not be aware that we are measuring the properties of individual photons. However, the fact that exactly the same results are obtained by measuring a very weak beam for a long time indicates that the basic phenomenon involved is the detection of individual particles.

To summarize: The quantum state refers to individual particles. Therefore it is incorrect, for example, to speak of a beam in the quantum state of x-polarization. On the other hand, when speaking of "a photon in a state of x-polarization," one always implies that the photon is one of a beam of photons verified to be in the quantum state of x-polarization.

The necessity for studying a large set of identically-prepared systems becomes even clearer when we look closely at the processes by which photons are detected. Photon counters used to detect photons in any state destroy the photons or at least change their energy. We have used a photomultiplier tube to detect photons. Photons impinging on the face of such a tube are absorbed, driving electrons out of the photosensitive surface. It is these electrons which are "multiplied" and counted. Use of the phototube unavoidably involves the destruction of the photons detected. And no observation on a beam of photons can be completed until the photons are detected. It is not necessary to destroy photons in order to detect them. One can, in principle, design a device in which the recoil of mirror with very little mass is used to signal the arrival of photons. Photons are thus detected without destroying them. However, even in this case, the process of reflection changes the photons: it changes their energy. Photons reflected from a recoiling mirror have reduced energy: they are in a new quantum state. The only way to avoid this change of photon energy is to use a large mirror that recoils a negligible amount. But then the recoil of the mirror cannot be used to detect the photons. Apparently a change of state inevitably accompanies detection. Diligent study has failed to produce a single method of detection which will not--at the very least--change the state of the photon or its energy in an

part of state spec.

unpredictable manner.

In classical physics we are accustomed to making repeated observations on a single (large) object. The situation is entirely different for the systems studied in quantum physics. Here the process of making an observation unavoidably alters the system being observed. At most one observation may be carried out on each system. How then can one study such systems at all? Only by employing a continuous production line on which systems--in this case photons--are prepared in a manner which is identical as far as we can tell. These "identical" systems are then subjected to different observations in which they are unavoidably altered or destroyed. Comparison of the results of different observations can then be used to describe the systems so produced. In each of the experiments described in this chapter the verification of a probability involves the comparison of intensities of two different beams. Two independent observations are carried out on different groups of photons. This is one more reason why statements as to what will happen to a particular photon are meaningless. Rather the predictions and results are always expressed in terms of relative average counts for different experimental arrangements.

An objection can be raised against the probabilistic interpretaion. Return once more to Experiment 4(Section 3 and Figure 4). Suppose for concreteness that the transmission axis of the first polarizer, A, is parallel to the x axis, whereas the transmission axis of the second polarizer, B, is oriented at some other direction (but not perpendicular to the x axis). Now, the beam incident on B is composed of x-polarized photons: allegedly identical particles identically prepared. Yet not all of these incident particles emerge from B; some are transmitted by polarizer B whereas others are absorbed. How can it be that so-called identical particles do not behave identically when subjected to identical treatment?

The term identity is useful in the sense that there do exist experiments for which a given outcome will always happen--for which the probability is unity. It is just such experiments that we use to define quantum states. Particles that are in a given quantum state witness that fact by behaving identically when tested for that quantum state. The search for quantum states is just the search for experiments that yield such identical behavior. In many other experiments, even particles in a given quantum state do not behave identically. But if we know the quantum state, we can always make definite statements regarding the relative probabilities of the possible outcomes of an experiment.

6. The calcite analyzer.

The experiments with linear polarizer, in terms of which we have defined linear polarization quantum states, are simple to carry out. However, we have had to assume properties that are only approximately realized by the actual material. Commercial--as opposed to ideal--linear polarizer absorbes a considerable fraction (typically 10 to 20 percent) of an incident beam linearly polarized parallel to the transmission axis of the polarizer. Furthermore, even ideal polarizer has the disadvantage that it absorbs part of an arbitrary incident beam (the part polarized perpendicularly to the transmission axis.) The absorbed photons are not available for further observation, and represent a loss of some of the energy and information carried by the incident beam.

All the functions of linear polarizer are exhibited by a different device called an analyzer. A linear polarization analyzer is a device that splits an incident beam into two physically separated linearly polarized beams without absorbing any of the light. Each separated beam can be studied individually. An analyzer is easily converted to a polarizer by blocking one of the output beams and allowing the other to proceed alone; the photons that emerge are then in one state of polarization only.

A linear polarization analyzer with the required properties can be constructed from a birefringent crystal, such as calcite. The phenomenon of birefringence is described in some detail by Shurcliff and Ballard (op. cit.) Chapter 5; Box 1 contains a brief summary.

BOX 1. The Birefringence of Calcite

When a beam of light passes obliquely across the plane boundary between air and glass the beam is refracted (bent) (Figure 9a). The refraction is governed by Snell's law (Eq. 5). A single index of refraction summarizes the refractive properties of glass, which are the same in all directions; glass is an isotropic dielectric.

Some crystals are more complicated. When a light beam is incident on a crystal of calcite (chemical formula $CaO \cdot CO_2$) there are in general two refracted rays. (Figure 9b.) One of these, called the "ordinary" ray, obeys Snell's law with a well-defined index of refraction; the other, called the "extraordinary" ray, has many unusual properties. For one thing, it is refracted even when light is incident normally on the surface (Figure 9c). Materials which exhibit this characteristic are called birefringent or doubly-refracting. Figure 10 shows the double images characteristic of objects observed through calcite.

Observations on the two beams separated by a calcite crystal show the following properties independent of the polarization of the incident beam,

i) each of the refracted beams in linearly polarized (calcite is said to be linearly birefringent),

ii) the polarization axes of the two refracted beams are perpendicular to each other; for a given orientation of the crystal and direction of incidence these polarization axes are also independent of the nature of the incident beam.

iii) absorption by the crystal is negligible.

We are not concerned here with the structural features of calcite that account for its unusual properties. Rather, we wish to exploit this birefringent property to construct an analyzer for

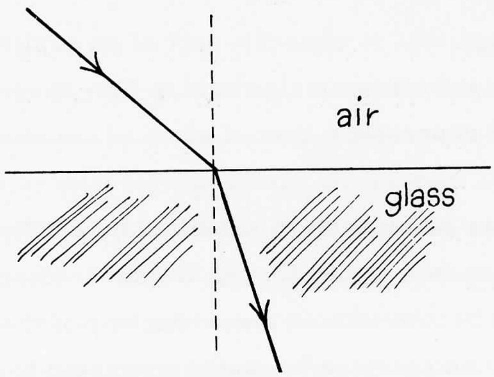


Figure 9a. Refraction of a ray of light when passing from air into glass.

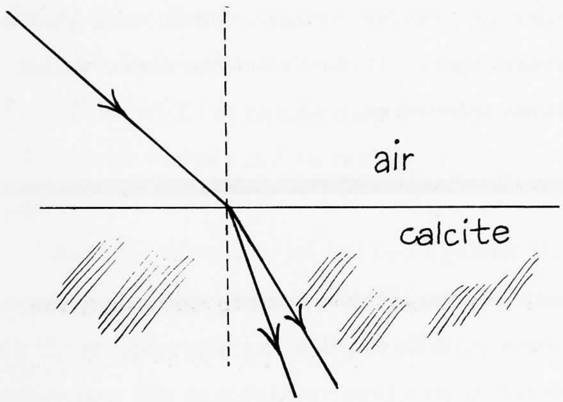


Figure 9b. Refraction of a ray of light into two beams when passing from air into calcite.

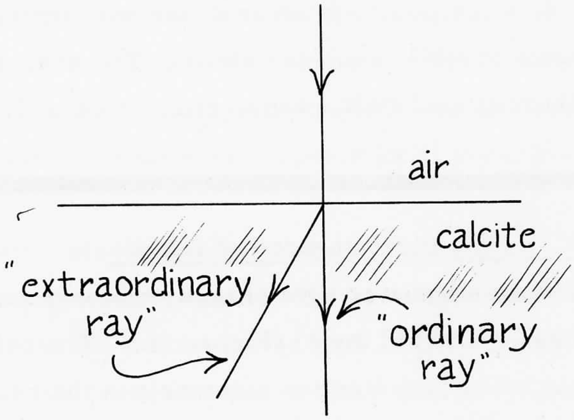


Figure 9c. The extraordinary ray can be refracted in calcite even when the incident beam is normal to the surface.

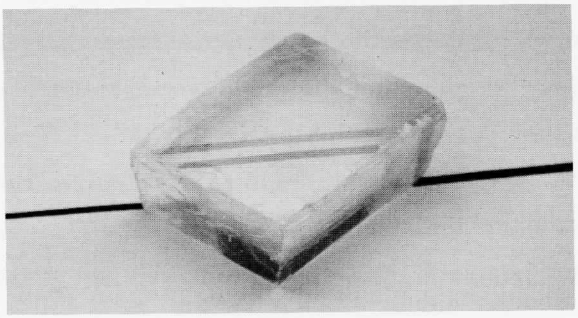


Figure 10. Photograph of a piece of natural calcite showing double images due to birefringence.

polarized light. Even a modest acquaintance with the properties of calcite enables us to place an order for a particular crystallographic cut with the properties shown schematically in Figure 11. Light incident normally on the crystal is split into an ordinary and an extraordinary ray; the extraordinary ray is refracted again on leaving the crystal and emerges traveling in the initial direction once again, but displaced laterally from the ordinary ray. The device with these properties operates as a linear polarization analyzer.

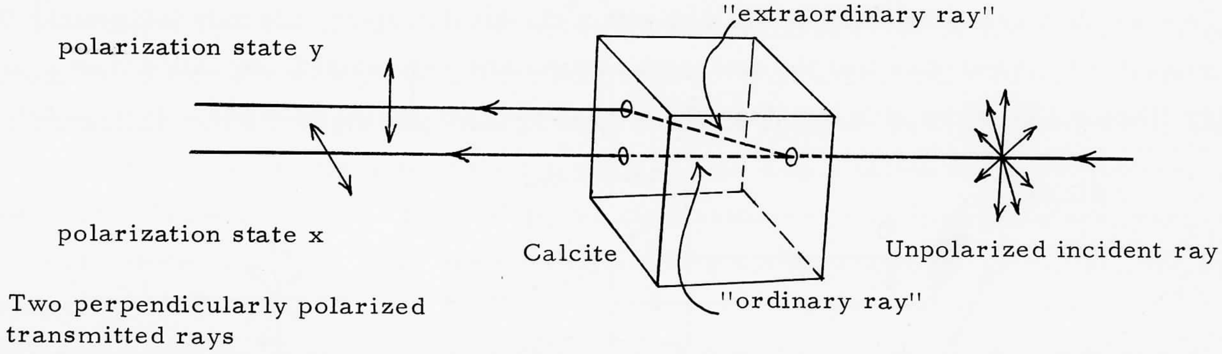


Figure 11. Schematic diagram of a calcite analyzer. In this orientation the device is called an xy analyzer.

As explained in the box, the arrangement sketched in Figure 11 analyzes an arbitrary incident beam of photons into two beams that are linearly polarized along mutually perpendicular axes. It may therefore be called a linear polarization analyzer. Let the z axis point along the direction of the input beam. Then the polarization axes of the output beams are in the xy plane. When the device is so oriented that the emergent beams are in polarization states x and y (Figure 11), we call it an xy polarization analyzer, or simply an xy analyzer. When the same device is rotated about the z axis to a new orientation, the output beams are in a different pair of polarization states which we designate as x' and y' (Figure 12). The device can then be called an x'y' analyzer. The labels x' and y' will in general designate an arbitrary pair of perpendicular axes in the xy plane, making an angle θ with the x and y axes respectively. By definition, if the beam incident on an x'y' analyzer consists of photons in the polarization state

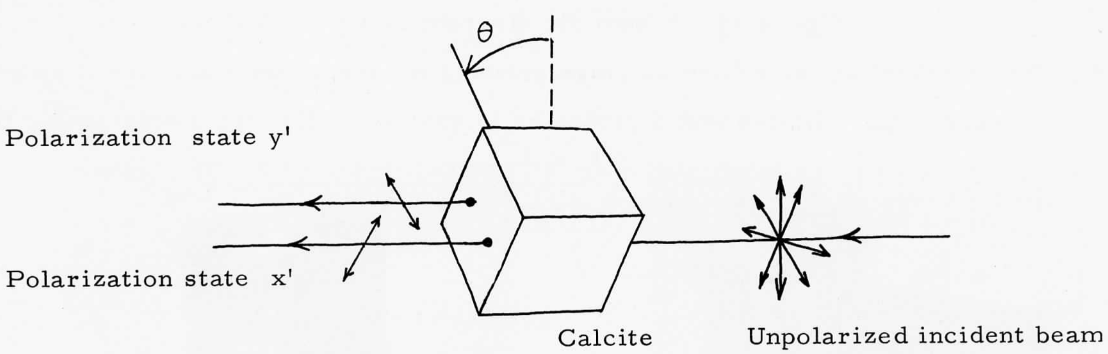


Figure 12. An x'y' analyzer.

x', the entire output will be found in the x' channel.

Let a stop be inserted in one of the output channels of a linear polarization analyzer. Photons in the remaining output channel of the analyzer are then in a known linear polarization state. We call this analyzer with only one channel open a linear polarization projector. The name comes from an analogy with the classical analysis of wave optics, in which the electric field of the output beam is the vector projection of the electric field of the input beam (Figure 5). (But we shall see later that this analogy with a classical projection is only incidental.) When the projector is oriented so that the emergent photons are x-polarized, we called it an x-projector (Figure 13). If we stop the other channel instead, the device becomes a y-projector

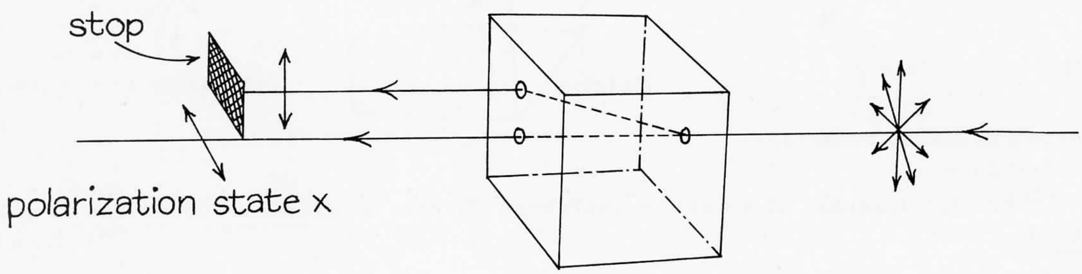


Figure 13. Schematic diagram of an x-projector.

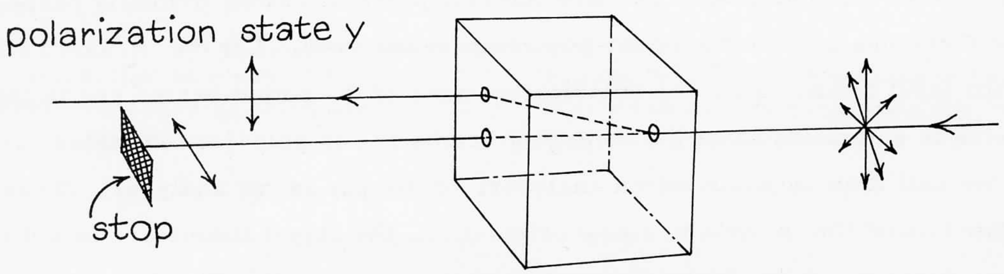
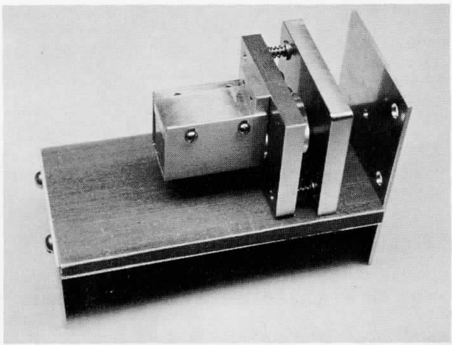


Figure 14. Schematic diagram of a y-projector.

(Figure 14). By re-orienting the device and manipulating the stops, we can make it either x'-projector or a y'-projector. Notice that a projector is automatically also a polarizer. By



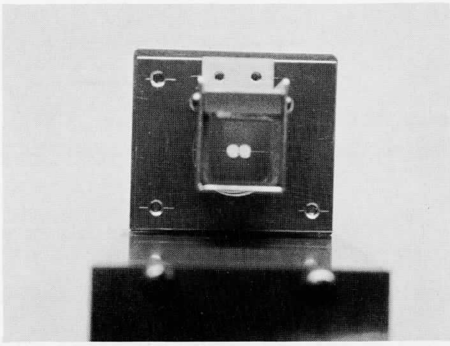


Figure 15. Photographs of a calcite analyzer. The photograph at the right shows the two separate beams (separated horizontally in this particular analyzer).

suitably orienting the device we can use it to produce a beam of photons in any desired state of linear polarization.

According to the definitions of the devices, a sheet of linear polarizer can serve as a projector but not as an analyzer. In this respect calcite is more useful than linear polarizing sheet. The other advantage of calcite is that an actual device which closely approximates the properties of an ideal analyzer can be constructed relatively easily. Figure 15 shows two photographs of such a device. A piece of optical quality calcite with polished surfaces and anti-reflection coatings transmits almost all the light incident upon it. This means that, on the average, (almost) the same number of photons emerge from a calcite analyzer as are incident on it, as can be verified by comparing the sum of the intensities of the output beams to the intensity of the input beam. In the discussions that follow, we shall assume that all the calcite analyzers employed are ideal--i.e., that there are no photon losses due to reflection, scattering, and absorption.

We now consider an experiment which is a generalization of experiment 4 of Section 3 using a calcite analyzer instead of the linear polarizer (Figure 16). For concreteness, let the incident beam be y-polarized, and measure the intensities of the beams emerging in the x' and y' chan-

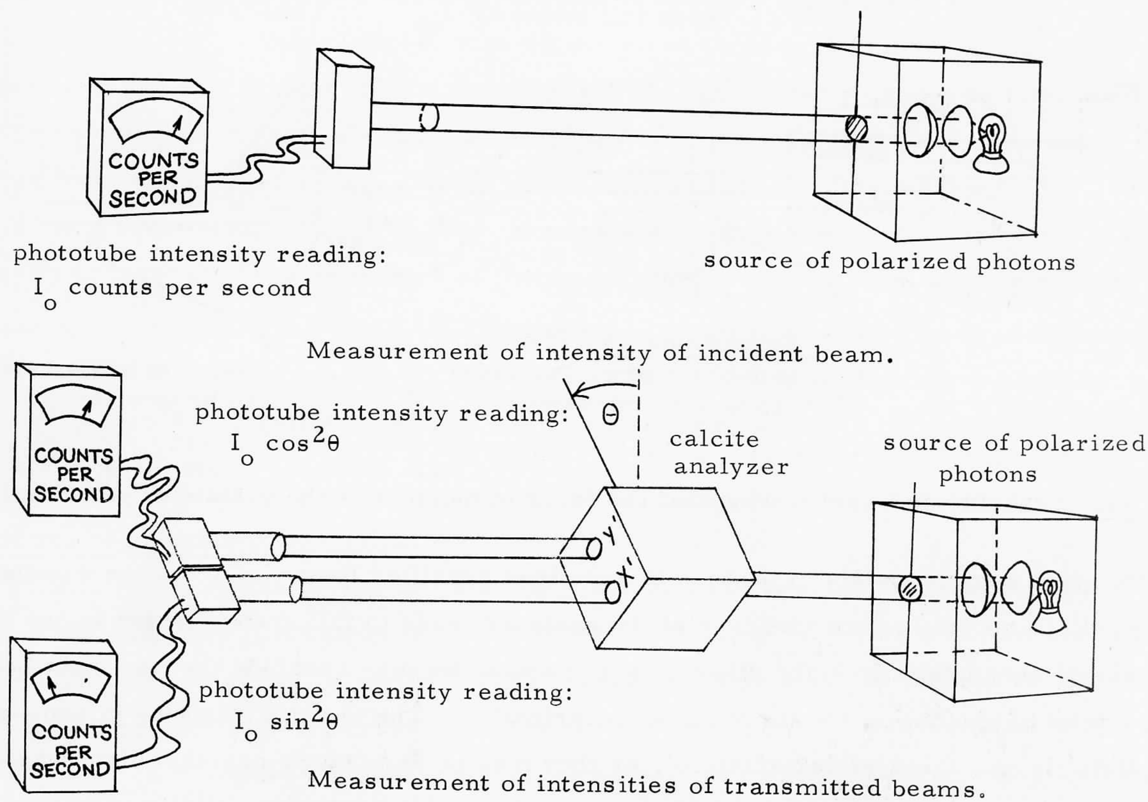


Figure 16. Experiment illustrating the effect of an ideal linear analyzer on a linearly polarized beam. The two separated beams together carry, on the average, the same number of photons as the incident beam.

nels as a function of the angle θ between the y axis and the y' axis. The intensities of the emerging beams are

$$\begin{aligned} I(y') &= I_o \cos^2\theta \\ I(x') &= I_o \sin^2\theta \end{aligned} \tag{7}$$

where I_o is the intensity of the incident beam. We thus confirm the values of the probabilities obtained first from our analysis of Experiment 4. This similarity between results of experiments employing structurally quite different devices suggests that the experiments indeed determine something about the properties of linearly polarized photons, and not merely about sheet polarizer or calcite.

The calcite analyzer provides an alternative test to determine whether the photons in a given beam are in a linear polarization quantum state, a test more practical than the one given at the beginning of Section 4: The photons in a beam of intensity I_o are in the state of y-polarization if and only if the insertion of an xy analyzer in the path of the beam results in intensity I_o in the y output channel and zero intensity in the x output channel. This verification can be repeated on the same beam as often as desired (Figure 17). One can of course, similarly test

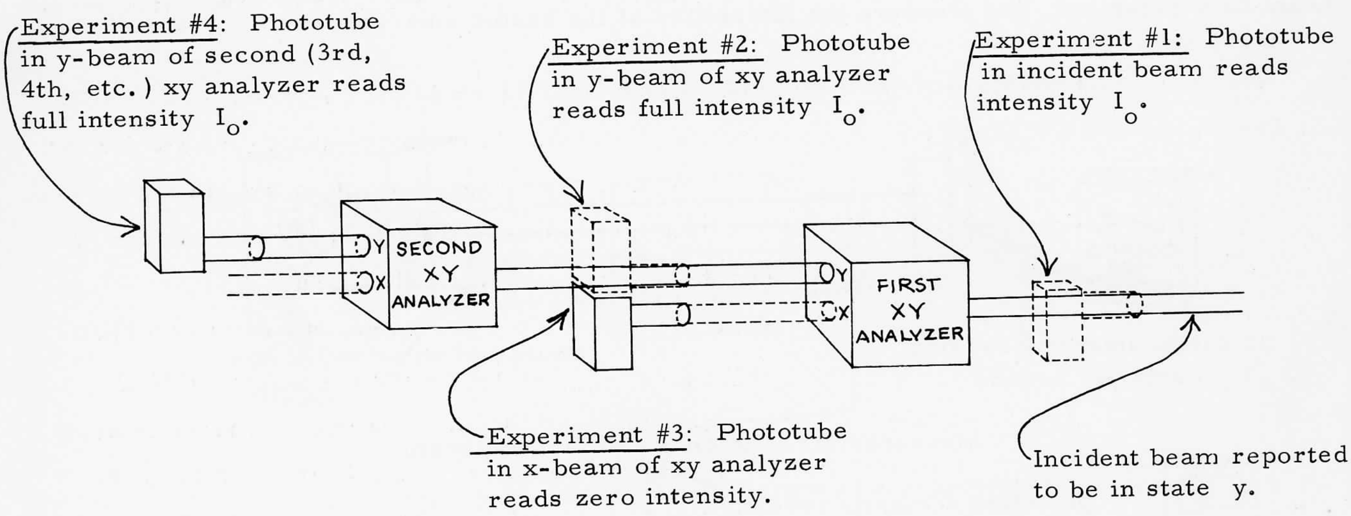


Figure 17. Experimental verification that the incident beam is in the y-state of polarization.

whether the given beam consists of photons in any other specified linear polarization quantum state, say y'. If no orientation whatever of the analyzer leads to full transmission in one output channel and zero intensity in the other output channel, we may conclude that the input beam does not consist of photons in a state of linear polarization. The photons might be in some other quantum state (e.g., circular polarization), or they may be in no single quantum state whateven. The experiments necessary to decide between these alternative possibilities will be discussed later.

When photon states are defined by means of the calcite analyzer, the labels x, y, x', etc., refer to both of the following:

i) the orientation of the analyzer,

ii) the particular output channel from which the beam in question has emerged.

As we have already remarked, the connection with the direction of the electric field in the classical wave description of the same beam is only incidental. For other quantum states, to be defined later, even this incidental connection is absent; the state is labeled by the experiments that define it.

We may now give a definition of an analyzer that will apply to quantum states in general. An analyzer for a given physical system is a device with one input channel and two or more output channels, such that:

1. if any beam whatever (as long as it consists of the particles in question) enters the input channel, the sum of the intensities of the output beams in all channels equals the intensity of the input beam;

2. if the beam in one particular output channel (say channel j) of the analyzer enters a second analyzer identical to the first in all respects, then the output of the second analyzer will be found entirely in the same channel j. This property also holds regardless of what beam is incident on the first analyzer.

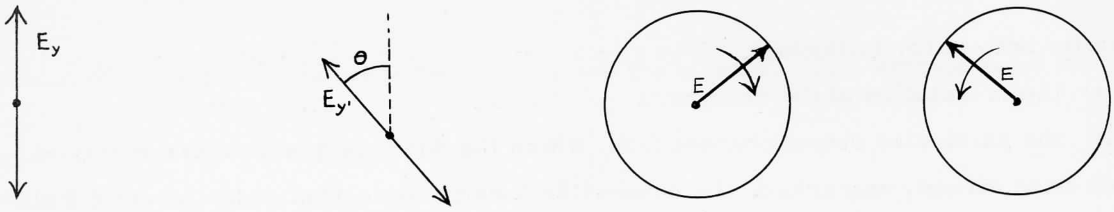
If we can construct for a given type of particle or more complicated system, a device that exhibits properties 1 and 2 above, then we are on our way to defining quantum states for that particle or system. It must still be verified that the output beams cannot be further subdivided.

A projector for a particular quantum mechanical state (say state j) of any system can be formed from an analyzer by blocking all output channels but the single channel j.

When a beam verified to be in state k passes through a j-projector, the ratio of output intensity to input intensity is defined as the projection probability from state k to state j. For example, referring to Figure 16, the projection probability from photon polarization state y to polarization state y' is $\cos^2\theta$. Similarly, the projection probability from state y to state x' is $\sin^2\theta$. Questions: What is the projection probability from state x to state x'? from state x to state y'? from state y to state x?

7. States of circular polarization.

Photon states of linear polarization, on which we have thus far concentrated our attention, are not the only ones that can be defined. In this section we introduce a second type of photon states--circular polarization states. These quantum states are, like the states of linear polarization, defined by re-interpreting in terms of the photon picture a phenomenon of classical wave optics. A classical circularly polarized wave is one whose electric vector is constant in magnitude and, at any fixed point, traces out a circle at a uniform angular velocity. (The magnetic vector does likewise.) If the electric vector rotates clockwise as observed looking toward the source (Figure 18c) the wave is called right circularly polarized; if the electric



a) linear y polarization   b) linear y' polarization   c) Right-circular polarization   d) Left-circular polarization

Figure 18. Linear and circular forms of polarization as interpreted in terms of a classical electric vector. In each case the diagrams are drawn as an observer would determine the fields as he looks toward the source (beam propagating perpendicularly out of the page).

vector rotates counterclockwise (Figure 18d) the wave is called left circularly polarized.* For a wave propagating in the positive z direction the electric fields of the two types of circular polarization may be written as follows:

right circular polarization:

$$E_x = -A \sin(kz - \omega t + \phi)$$
$$E_y = A \cos(kz - \omega t + \phi) \qquad (8)$$

Left circular polarization:

$$E_x = A \sin(kz - \omega t + \phi)$$
$$E_y = A \cos(kz - \omega t + \phi) \qquad (9)$$

Here ϕ is an arbitrary phase factor which determines the direction of the electric field at a specific time and place; for example, if $\phi = 0$ the field at $t = 0$, $z = 0$, points in the positive y direction both for the right circularly polarized wave (8) and for the left circularly polarized wave (9). It is clear from equations (8) and (9) that the magnitude of the field is constant in both space and time. The intensity of the wave is proportional to A^2.

Notice that a circularly polarized beam can be considered as a superposition of two beams linearly polarized along mutually perpendicular axes (e.g., x and y), with a 90° phase difference between the vibrations. This property forms the basis for one simple method of producing a circularly polarized beam, using linear polarizer and a quarter wave plate. Box 2 summarizes the properties of a quarter wave plate and describes a circular polarizer. A similar procedure can be used to detect circularly polarized beams (see the exercises).

BOX 2. The Quarter Wave Plate

In Box 1 we have examined some of the properties of a birefringent calcite crystal, and described one way of cutting such a crystal which makes possible the construction of a linear polarization analyzer. A different crystallographic cut of the same material leads to quite different transmission properties that are useful in other applications. A cut can be found such that for normal incidence, both the ordinary and extraordinary rays continue unrefracted, but with different velocities of propagation. The polarizations of the two rays are perpendicular,

* Caution: in some books the definition of left- and right-circularly polarized light are reversed (i.e., one looks along the beam away from the source instead of toward the source).

as in the previous application. The extraordinary ray travels faster. We therefore call its polarization axis the fast axis of the crystal and the polarization axis of the ordinary ray the slow axis. A crystal cut in this manner makes possible the construction of extremely useful devices known as quarter wave plates and half-wave plates. We shall describe these devices in classical language entirely.

Consider a linearly polarized ray normally incident on a crystal cut in the manner referred to above. The component of the electric field of this ray along the "ordinary" axis emerges from the crystal retarded in time relative to the component along the "extraordinary" axis--the magnitude of the time retardation depending upon the thickness of the slab. Corresponding to this relative retardation in time is a relative difference in phase between the emerging rays--the magnitude of the relative phase difference depending upon the thickness of the slab. Thus the net effect of the slab of calcite is to change the relative phase of the two perpendicularly polarized components. By a correct choice of the thickness of the calcite slabs, one can obtain any desired value for this phase shift. The two most-used thicknesses are the so-called "half-wave plate"--which produces a phase angle shift of π; and the so-called "quarter-wave plate" --which produces a phase angle shift of $\pi/2$. We will use here only the quarter-wave plate. It must be emphasized that one constructs a quarter-wave plate for light of a particular frequency or narrow range of frequencies. Only for this range of frequencies does the particular relative time delay between the two coincident perpendicularly polarized beams correspond to a relative phase angle shift nearly $\pi/2$. For light of a different frequency the relative phase shift will be other than $\pi/2$. Quarter-wave plates perform the tasks described below over a much narrower range of frequencies than sheet polarizer and calcite separators perform their functions. But this is not a serious drawback since, if necessary, all the experiments can be carried out with monochromatic beams.

Let a quarter wave plate be oriented so that its fast and slow axes point in the y and x directions, respectively, and let an incident beam be linearly polarized at 45° (Figure 19).

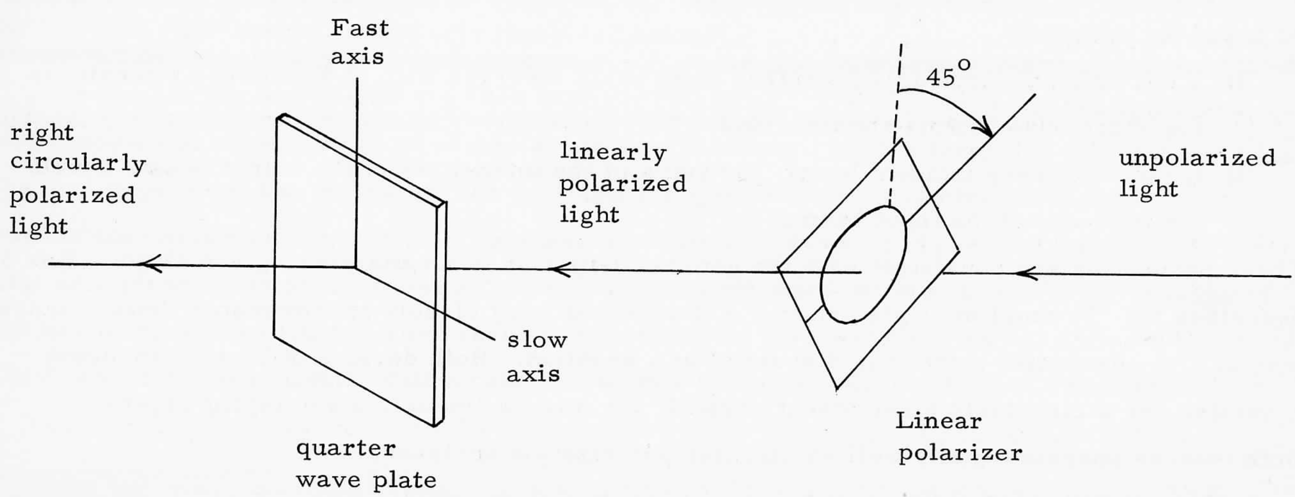


Figure 19. Production of right-circularly polarized light using linear polarizer and a quarter-wave plate.

The electric field of this beam can be written as

$$\begin{aligned} E_y &= A\cos(kz - \omega t) \\ E_x &= A\cos(kz - \omega t) \end{aligned} \qquad \text{input beam} \qquad (10)$$

After the beam has passed through the quarter wave plate, the extraordinary ray has acquired some extra phase ϕ, because its velocity in the crystal in not equal to c. But the ordinary ray, which travels slower, has acquired an additional phase $\pi/2$. Therefore we can write the electric field of the emergent beam as

$$\begin{aligned} E_y &= A\cos(kz - \omega t + \phi) \\ E_x &= A\cos(kz - \omega t + \phi + \pi/2) \end{aligned} \qquad \text{output beam} \qquad (11)$$

This outgoing wave is right-circularly polarized. Consequently the quarter wave plate can be used to convert a particular linearly polarized wave to circular polarization. In conjunction with a linear polarizer, it can be used to convert an unpolarized beam into a circularly polarized beam. One kind of right circular polarizer available commercially consists of a sheet of quarter wave plate bonded to a sheet of linear polarizer with properly oriented transmission axis. A different relative orientation of the transmission axis yields a left circular polarizer. This and further details are left to the exercises.

Very useful for our purposes is the circular polarization analyzer (Figure 20). An ideal

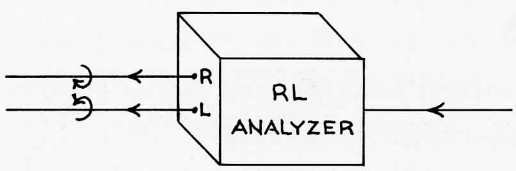


Figure 20. Idealized RL analyzer.

circular polarization analyzer is a device with one input channel and two output channels, labeled L and R, such that

i) a left or right circularly polarized input beam emerges with undiminished intensity in the appropriate output channel, and

ii) for an arbitrary incident beam, the sums of the intensities of the output beams equals the intensity of the input beam.

These properties are consistent with the general definition of an analyzer, given above. Box 3 describes the "Fresnel multiple prism," a device that very closely approximates these requirements. An alternative method is discussed in a problem. Both devices utilize birefringent crystals, one a circularly birefringent crystal, the other a linearly birefringent crystal. Both devices operate equally well as circular polarization analyzers.

BOX 3. Design of an RL analyzer

The linear birefringence of crystals such as calcite has been used in the construction of the linear polarization analyzer. There exist also crystals that are circularly birefringent; such crystals can split an incident beam into two beams, one of which is left-circularly polarized and the other right-circularly polarized. Quartz is a common example of a circularly birefringent material. Right and left circularly polarized light follow different paths through a piece of quartz properly cut and oriented with respect to the incident beam. From such a piece of quartz one can construct an "RL analyzer" as idealized in Figure 20.

In practice the angle between the separated right- and left-circularly polarized beams in quartz is very small. Therefore adequate separation between beams can be achieved only by using a piece of quartz which is long in the beam direction. Alternatively there is a way to use several prisms of different crystallographic forms of quartz cemented together, as shown in Figure 21.* This device makes use of the different speeds of R and L light in quartz to achieve different refractions at prism interfaces. A third model of the RL analyzer may be constructed from three pieces of calcite, as outlined in the exercises. In summary one can say that it is technically more difficult to manufacture a circular polarization analyzer than it is to manufacture a linear polarization analyzer. This technical difficulty in no way affects the similarity in principle between the RL separation achieved by a circular polarization analyzer and the xy separation achieved by a linear polarization analyzer.

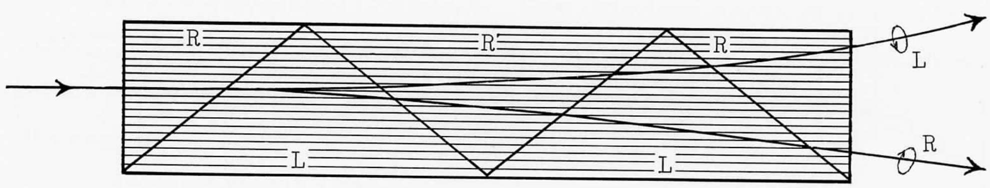


Figure 21. "Fresnel multiple prism" composed of alternate prisms of "right-handed quartz" (labeled R) and "left-handed quartz" (Labeled L). This device analyzes a beam incident from the left into right and left circularly polarized beams.

The preceding discussion has been entirely classical. The Fresnel multiple prism is a well-known and useful device of classical optics. However, when interpreted on the basis of the photon picture, the device serves to define the quantum states of circular polarization, just as the calcite analyzer of Section 6 served to define states of linear polarization. The photons in a given beam of intensity I_o are in the quantum state R if and only if the passage of the beam through an RL analyzer results in intensity I_o in the R channel and zero intensity in the L channel. Verification that a beam is in state L proceeds similarly. This definition ful-

* Francis A. Jenkins and Harvey E. White, Fundamentals of Optics, McGraw-Hill Book Co., Inc., New York, 1957, page 581.

fills the requirements for a quantum state set forth in Chapter 1 and discussed in detail earlier in this chapter for states of linear polarization.

If an RL analyzer is rotated about an axis along the direction of the incident light, its effect remains unchanged: the output beams are still right- and left-circularly polarized. There are no states R' or L', distinct from R and L. In this respect circular polarization differs from linear polarization.

The RL analyzer can evidently be turned into an R-projector or an L-projector by blocking the appropriate output channel. Notice that the term projector as used for circular polarization has a purely quantum mechanical meaning; there is no analogous classical projection of a vector onto some axis, as is the case for a linear polarization projector. The projectors also constitute polarizers--they provide a means for producing circularly polarized light. They can be used instead of the sheet circular polarizer described in Box 2 (which is a polarizer but not a projector: see the exercises).

In Section 6, we defined the projection probability from one quantum state to another, and considered the projection probabilities between different linear polarization states. We can now inquire as to the projection probability from a state of linear polarization to one of circular polarization, or vice versa. The experiments which measure these probabilities are straightforward: to determine the projection probability from state R to state y, for example, one merely sends a beam of R-polarized photons into a y-projector (Figure 22), and measures the relative intensities of input and output beams. The results of all such experiments can be pre-

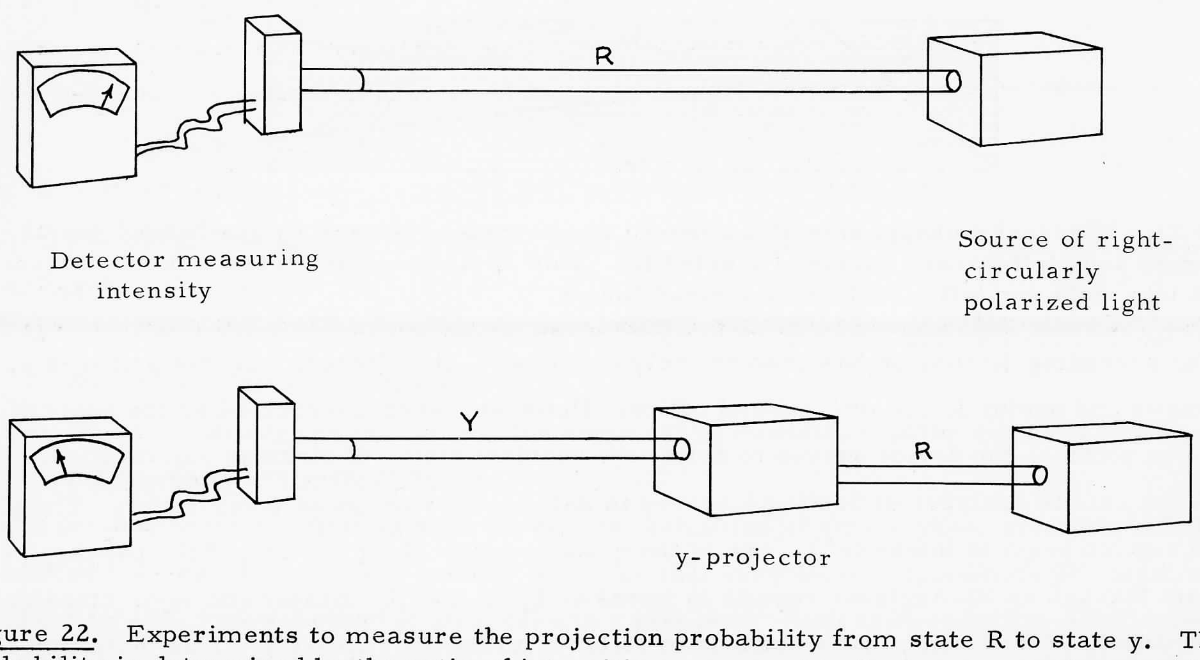


Figure 22. Experiments to measure the projection probability from state R to state y. This probability is determined by the ratio of intensities measured in the lower experiment and the upper one. Classical arguments predict correctly that this projection probability should be one-half (see text).

dicted on the basis of the classical wave picture. As we have already remarked, a classical right-circularly polarized wave can be considered as the superposition of x- and y-linearly polarized waves with equal amplitudes. Therefore, in the example under consideration, we would expect to find equal intensities in the x and y output channels (if they were both open). The classical argument therefore suggests that the projection probability from state R to state y is 1/2; in fact, similar arguments suggests that each of the eight possible projection probabilities (x to R, x to L, R to x, L to x, etc.) has the value 1/2. This prediction is confirmed when the experiments are performed. In the analysis of photon polarization, all predictions of this kind based on the classical wave picture turn out to be correct. When we consider phenomena for which no classical picture exists, as in Chapter 7, we will not be able to predict the projection probabilities, but must take them from experiment.

8. Orthogonality and completeness

The polarization states of photons exhibit several formal properties that simplify and summarize our knowledge of them. Among these properties are orthogonality and completeness. The quantum states of other particles and systems also exhibit these same properties.

Orthogonality. When the projection probability from some quantum state j to some other state k, is zero, we say that state k is orthogonal to state j. According to this definition the linear polarization state x is evidently orthogonal to state y. Also state y is orthogonal to state x. Likewise state R is orthogonal to state L, and vice versa. These examples illustrate the general rule that orthogonality is a reflexive property: if (1) state k is orthogonal to state j, then (2) state j is orthogonal to state k. One may thus speak of two states being orthogonal to each other. It is by no means obvious that statement (1) implies statement (2), since the two statements refer to entirely separate experiments. However, the rule is verified experimentally in all cases.

In plane geometry orthogonal is a synonym for perpendicular. And indeed, in the case of linear polarization, two quantum states are orthogonal if the polarization axes of the corresponding classical waves are perpendicular. This is in fact the origin of the term orthogonal. However, no such classical analogy exists for the statement that states R and L are orthogonal. The quantum mechanical concept of orthogonality is defined entirely in terms of experiments with analyzers and projectors (Figure 23).

Completeness When a photon beam of arbitrary polarization enters an xy analyzer, all outgoing photons are either x-polarized or y-polarized, and no photons are absorbed. In this sense, (to be developed in detail in later chapters) two states x and y are "sufficient" to describe an arbitrary beam. In mathematical language, the sum of the projection probabilities from any initial state to the states x and y is unity. When this is the case, one says that the

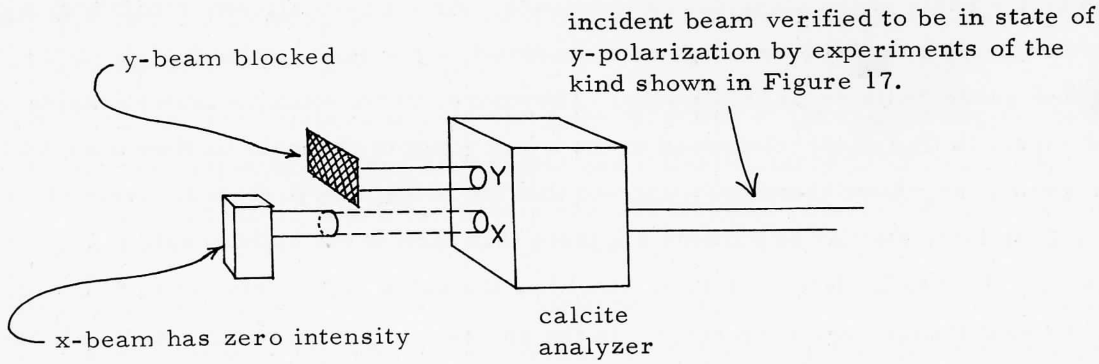


Figure 23. Verification that two states are orthogonal. Two states are orthogonal by definition when an incident beam known to be in one state has zero projection probability into the other state when tested by an analyzer or a projector.

states x and y constitute a complete set of photon polarization states. Alternatively, the states x and y provide the basis for a representation of photon polarization. There are, of course, many other complete sets. In fact, every analyzer defines a complete set. For example, states L and R, as well as states x' and y', constitute complete sets. But all the complete sets have one thing in common: for photon polarization each consists of exactly two states. Each analyzer has exactly two output channels; no third or fourth channel is needed. This fact reflects a fundamental property of photons. Later on we shall study systems for which a complete set consists of larger numbers (eventually, even an infinite number) of states.

Notice that in all the examples above, the two states that form a complete set are orthogonal to each other. Orthogonality, it turns out, is not a necessary condition for the members of a complete set (see the exercises). But it is always possible to find complete sets of states for any system that are orthogonal, and these are by far the most convenient to use. Notice also that the dimensionality of the representation (i.e., the number of states in a complete set) is the same as the maximum number of mutually orthogonal states that can be found. In the present example, there are many pairs of orthogonal states, but one cannot find three photon states that are mutually orthogonal. This remark has a geometric analog: in a space of n dimensions, at most n vectors can be mutually orthogonal. (i.e., "perpendicular."). The analogy between quantum states and vectors in an abstract space of n dimensions will be developed more fully in Chapter 6.

Table 1 contains a brief definition of the major quantum concepts that have been introduced in this chapter, together with specific examples for photon polarization to illustrate these concepts.

9. Elliptical polarization

Linear and circular polarization are not the only possible forms for a classical wave. The most general electric field for a coherent wave propagating in the positive z direction can be

Table 1. Some concepts and definitions of quantum mechanics

CONCEPT	GENERAL MEANING	EXAMPLE FOR POLARIZED PHOTONS
analyzer	An analyzer for a given physical system is a device with one input channel and two or more output channels such that: 1. If any beam whatever enters the input channel, the sum of the intensities of the output beams equals the intensity of the input beam. 2. If the beam in one particular output channel (say channel j) of the analyzer enters a second analyzer identical to the first in all respects, then the output of the second analyzer will be found entirely in the same channel j.	An xy analyzer made from calcite has one input channel and two output channels, x and y. Its properties closely approximate the following: 1. If any beam of photons enters the input channel, the sum of the intensities of the two output beams equals the intensity of the input beam. 2. If the output beam in channel y of the analyzer enters a second xy analyzer, then the output of the second analyzer is entirely in channel y (Figure 17).
projector	A projector for a particular physical system can be formed from an analyzer by blocking all output channels but one, say the channel j. Then the device is called a j-projector.	A y-projector can be constructed from an xy analyzer by blocking the x output channel.
determination of the state of a system	When the output of a j-projector is equal in intensity to the input, the input systems are said to be in state j.	When the output of a y-projector is equal in intensity to the input, the input photons are said to be in state y.
projection probability	Suppose an input beam is known to be in state k. (For example it may have come from a k-projector.) Let the beam now pass through a j-projector. The relative intensity (output intensity)/(input intensity) is called the projection probability from state k to state j. Experimental result: The projection probability from state k (input) to state j (output of a j-projector) has the same value as the projection probability from state j (input) to state k (output of a k-projector).	Experimental result: The projection probability from state y to state y' is $\cos^2\theta$. Here θ is the angle through which the y-projector must be rotated to make it a y'-projector. The general expression $\cos^2\theta$ implies that the "reverse" projection probability from state y' to state y also has the value $\cos^2\theta$. ($y \leftrightarrow y'$ implies $\theta \leftrightarrow -\theta$ and $\cos^2\theta$ is even.)
orthogonal states	Two states are orthogonal by definition if the projection probability from one state to the other is zero.	Polarization states x and y are orthogonal, since the output of a y-projector is zero when the input photons are known to be in state x, and vice versa.
complete set of states	The set of states defined by an analyzer is complete because, on the average, the number of particles emerging in the output channels is equal to the number entering the input channel.	The set of polarization states x and y is complete because, on the average, the number of photons emerging in the output channels of an xy analyzer is equal to the number entering the input channel.

written

$$E_y = E_{oy}\cos(kz - \omega t + \phi_1)$$
$$E_x = E_{ox}\cos(kz - \omega t + \phi_2) \tag{12}$$

The polarizations already considered are included in this general form: when ϕ_1 and ϕ_2 are equal, or differ by a multiple of π, eqs. (12) describe a linearly polarized wave. When $E_{ox} = E_{oy}$, and ϕ_1 and ϕ_2 differ by $\pi/2$ or $3\pi/2$, eqs. (12) describe circular polarization. For arbitrary values of the amplitudes and phases, the electric vector at any fixed point traces out an ellipse in the xy plane (Figure 24). Accordingly, such a wave is to be elliptically polarized.* The polarization is characterized by three parameters:

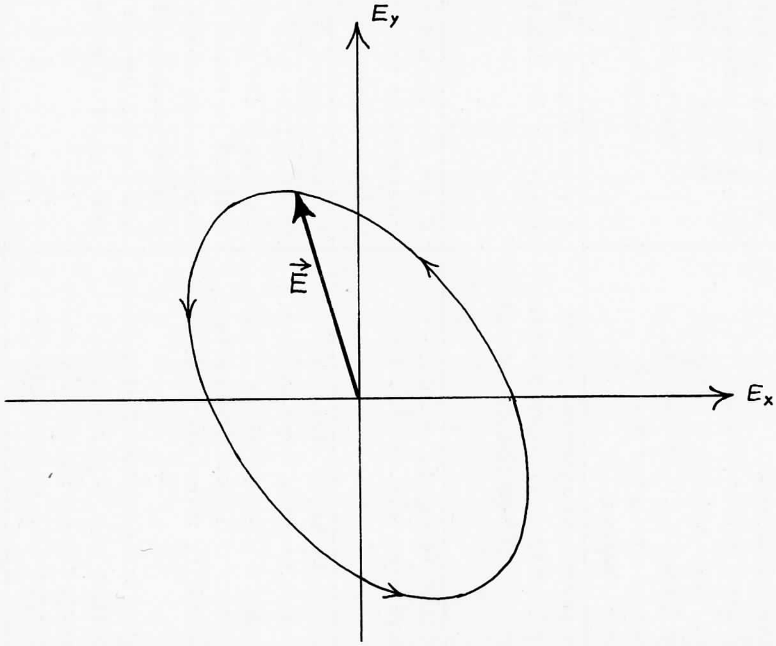


Figure 24. Elliptically polarized light. The ellipse shown corresponds to Eq. (12) with $E_{oy} = E_{ox}$, $\phi_1 = 0$, $\phi_2 = 4\pi/3$. The major axis then makes an angle of $\pi/4$ with the y-axis, the ratio of major axis to minor axis is $\sqrt{3}$, and the ellipse is traversed counterclockwise as one looks along the beam toward the source from a fixed point (beam emerging perpendicularly out of page).

i) the orientation of the major axis;

ii) the eccentricity of the ellipse.

iii) the sense in which the electric vector traces out the ellipse (clockwise or counterclockwise).

To each specific type of elliptical polarization, i.e., each distinct set of parameters i) to iii), there exists a quantum state for photons. To verify this assertion, one must construct appropriate analyzers and projectors; this can be done with the help of linear analyzers and quar-

---

* When one speaks specifically of an elliptically polarized beam it is usually implied that the beam is not in one of the special cases linear or circular polarization.

ter wave plates. For each state of elliptical polarization there is another state orthogonal to it, and together the two states constitute a complete set in terms of which any incident beam may be analyzed. Further details are left to the exercises.

10. Unpolarized light.

Not all light beams are polarized. If the light from an ordinary tungsten bulb is passed through a linear polarization analyzer, the output channels will be found to contain (very nearly) equal intensities for any orientation of the analyzer. The same result will be obtained if a circular or elliptical polarization analyzer is used. A beam which gives equal intensity in the output channels of every analyzer whatsoever is said to be unpolarized.

The classical wave interpretation of unpolarized light may be summarized as follows: the electric field at a given point, over a short enough time period, traces out some definite pattern. That is to say, every short "wave train" must have some definite polarization. But the wave train that passes a given point at one instant in general comes from a different part of the source than the wave train that passes at another instant; the polarizations of the two trains therefore need not be the same. In an unpolarized beam, the polarizations of the various wave trains that make up the beam are completely uncorrelated. If one observes such a beam over a period long enough to include many wave trains (as is the case in most experiments) the random assortment of polarizations will lead to equal intensities in the output channels of any analyzer.

How should unpolarized light be described in terms of the photon picture? In analogy to the classical interpretation above, one may say that each photon in an unpolarized beam is in some quantum state, but the beam consists of a random collection of photons in all possible states. This assertion, however, is not subject to direct verification. As we have already emphasized in Section 5, it is impossible to determine the quantum state of any single photon.

A beam can also be partially polarized. Such a beam gives unequal intensities in the output channels of some analyzers (not necessarily all), but never goes entirely into a single output channel of any analyzer.

CHAPTER 4. PROBABILITY AMPLITUDES

1. Introduction

In the preceding chapters we have defined a quantum state and examined some properties of the quantum states of a particular particle, the photon. We have defined the projection probability from one state to another as the fraction of incident particles in the first state that appear in the output beam of a projector for the second state.

In the present chapter we show that projection probabilities alone are not sufficient to describe the results of some additional experiments with photons. In order to describe these results we will be led to define a new quantity, the projection amplitude, a complex number whose absolute square equals the corresponding projection probability. The projection amplitude is one kind of probability amplitude. Every probability calculated using quantum physics is the absolute square of some probability amplitude. Complex probability amplitudes can describe interference experiments, in which two recombining beams cancel at some points to produce zero resultant intensity. In contrast, probabilities are fundamentally positive and as such do not provide a sufficient description of experiments in which such cancellation takes place.

2. Analyzer loop

The experiments we wish to consider involve an additional piece of equipment which we call an analyzer loop. An analyzer loop for photons is a two-part device, of which the first part is just an analyzer as defined in Chapter 3. The second part of the device is a "reversed" analyzer of the same type, which recombines the beams separated by the first analyzer in such a way as to reconstruct the original incident beam exactly. Achieving the proper recombination is no mean technical feat. It is not enough simply to superpose the separated beams geometrically. Each part of the cross sectional area of the original beam must be restored to the same position relative to the other parts, and care must be taken to insure that the optical paths traveled by the separated beams are equal or at most differ by an integral number of wavelengths. The optical accuracy required to construct a successful analyzer loop is comparable to that required to construct the most sophisticated optical equipment. In fact, the analyzer loop is actually a type of interferometer, as we shall shortly show. Schematic designs for analyzer loops are shown in Figures 1 and 2; Figures 3 and 4 are pictures of actual operating devices.

Following the nomenclature of the preceding chapter, we label an analyzer loop by the set of states associated with the analyzers which make up the device. Thus we speak of an xy analyzer loop or an x'y' analyzer loop (the same device with a different orientation) or an RL analyzer loop. All three of these devices behave identically as long as both channels remain

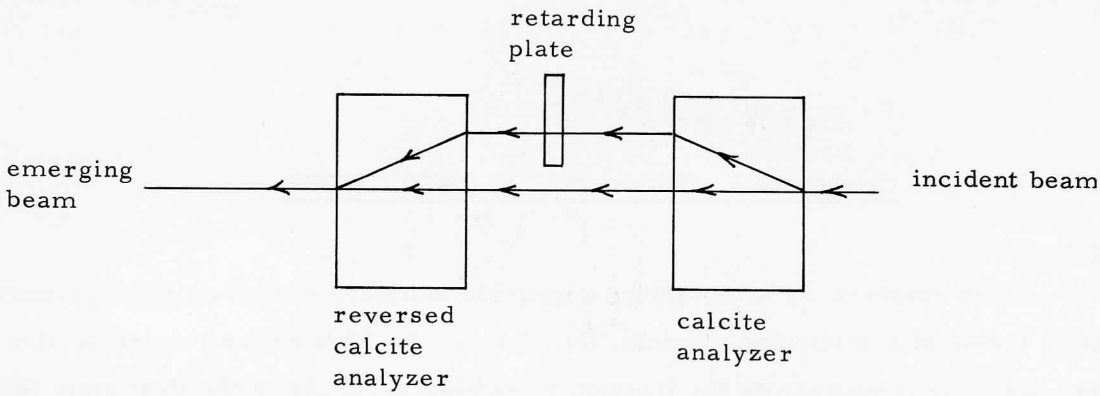


Figure 1. Diagram of analyzer loop constructed of two pieces of calcite and a retarding plate. The retarding plate is used to assure that the relative classical phase of the two beams on recombination is the same as at separation; it has no effect on the polarization.

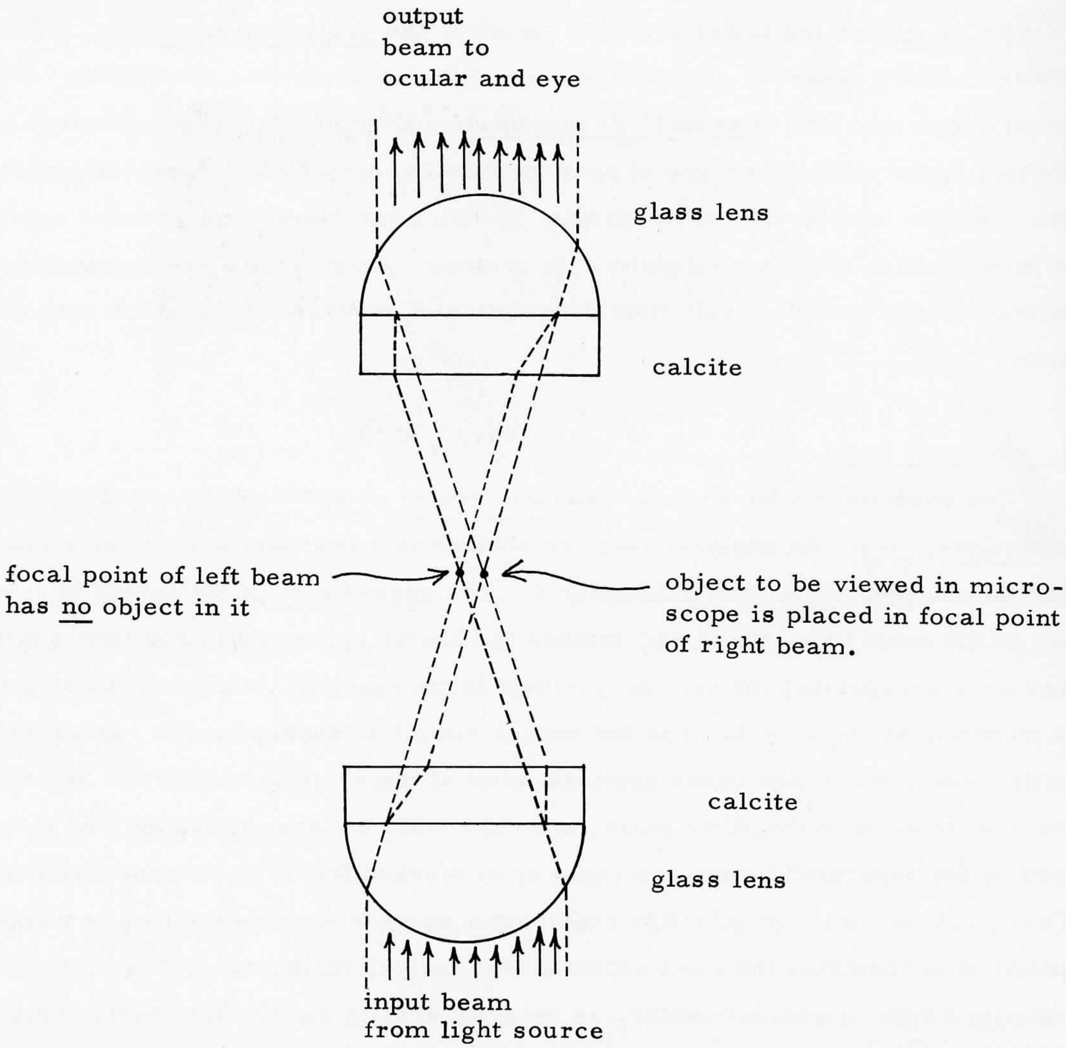


Figure 2. Diagram of an analyzer loop commercially available as a portion of one kind of interference microscope. When the loop is used in the microscope, the sample under observation is placed in one of the separated beams. Interference effects in the recombined beam allow determination of thickness and optical properties of the sample. The first published account of the construction of what we call an analyzer loop described a device similar to the one pictured here. (A. A. Lebedev, Revue d'Optique, 9, 385 (1930). See also John Strong, Concepts of Classical Optics, W. H. Freeman Co., San Francisco, 1958, p. 388.)

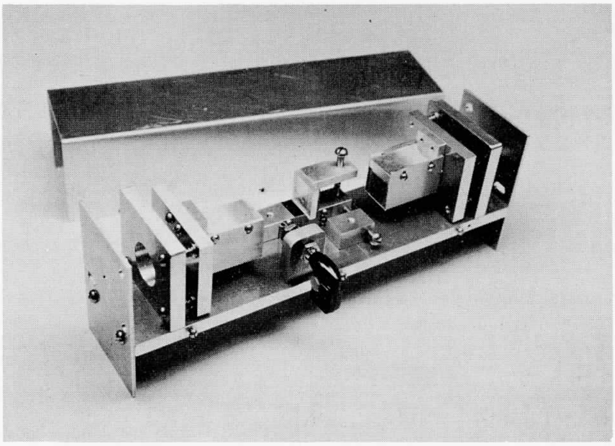


Figure 3. Picture of linear analyzer loop constructed after the design of Figure 1. Knob controls vanes (not shown) to stop one or the other of the separated beams when desired. Retarding plate and its mounting may also be seen in the center. This device was constructed by J. Burkhardt.

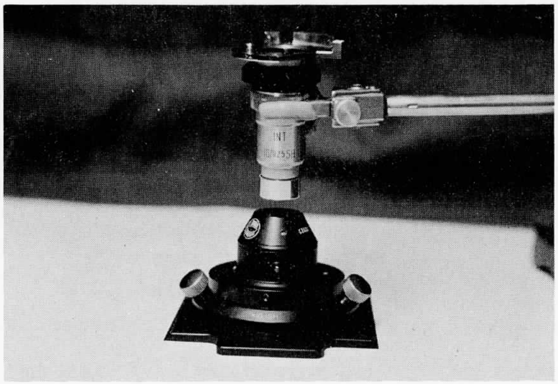


Figure 4. Picture of the portion of an interference microscope diagrammed in Fig. 2.

open: each device reconstructs all incident beams without change. Figure 5a to 5e illustrate simple experiments which verify that for a number of different inputs the analyzer loops are indeed operating as required. (Notice that the incident beam need not be in a state of x or y polarization, or in any state of linear polarization at all, in order to be transmitted unchanged by an open xy analyzer loop. See in particular Figure 5e.)

The concept analyzer loop is readily extended to quantum systems other than photons. In all cases we define an analyzer loop as a device which:

1. analyzes an incident beam into a number of physically separated beams, each in a single state. Taken together these states constitute a complete set of orthogonal states for the system in question. The analyzer loop then

2. recombines the separated beams in such a way as to reconstruct the incident beam exactly.

Inasmuch as the ideal analyzer loop has no observable effect on any beam, its usefulness at first sight appears questionable. Nonetheless the device has at least two important uses: First, by inserting stops or retarding plates in one of the channels, one can create more

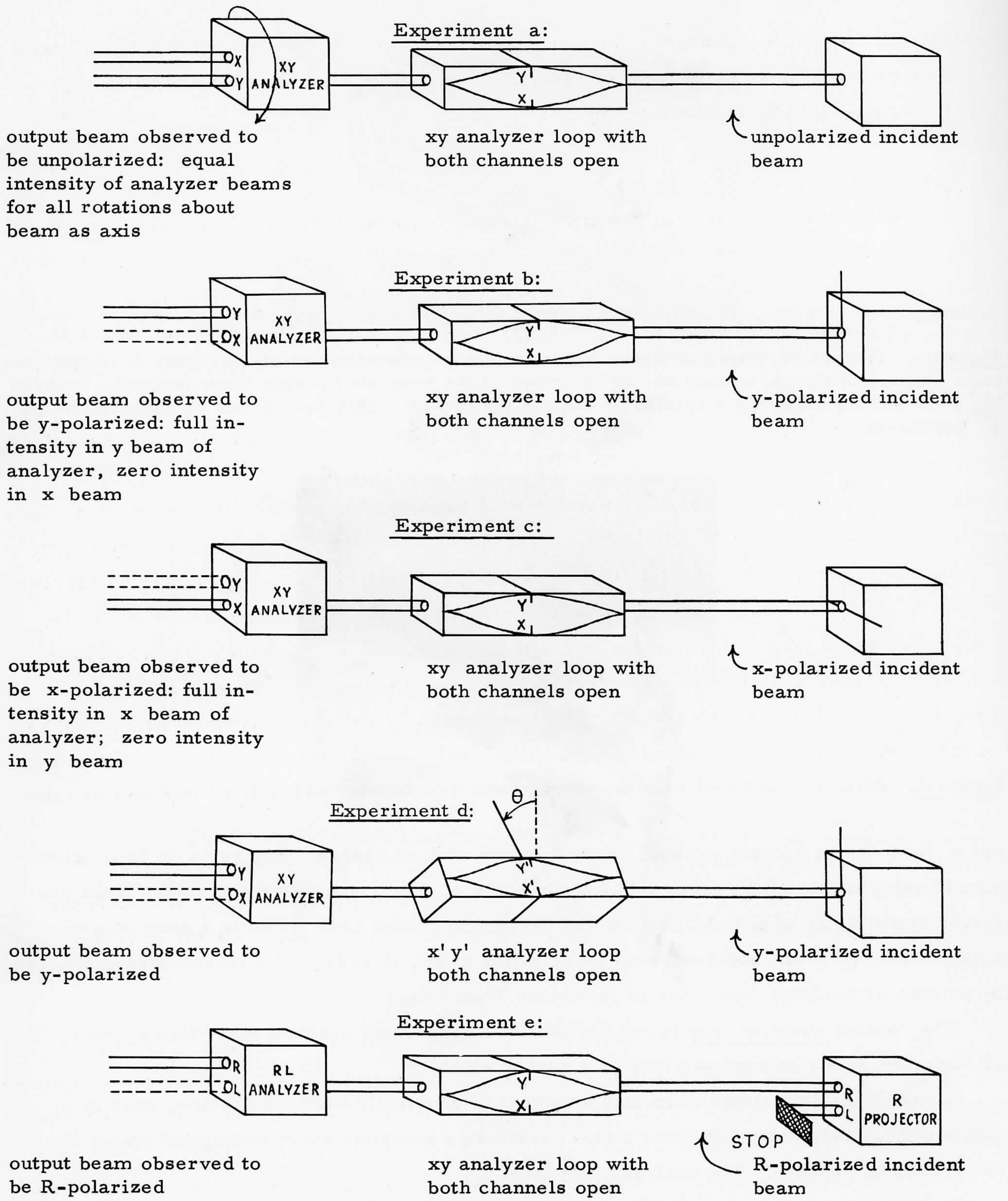


Figure 5. Experiments to demonstrate the fundamental property of an analyzer loop: any incident beam is transmitted unchanged.

general devices that do influence an incident beam in a prescribed manner. Second, by interpreting the operation of an analyzer loop in an appropriate manner one is led to the idea of probability amplitudes, as shown in the next section.

3. Paradox of the recombined beams

A simple extension of the experiment of Figure 5d demonstrates the inadequacy of projection probabilities to describe all polarization experiments with photons, and leads to the idea of amplitudes. The complete experiment consists of three parts, as diagrammed in Figure 6. A y-polarized beam passes through the x'y' analyzer loop A. For concreteness, let the angle θ between the y and y' axes be 30°. The output of A passes into the x projector B, and the intensity of the beam that emerges from B is measured.

In the first part of the experiment, the x' channel of the analyzer loop has been blocked (Figure 6a). When one interior channel of any analyzer loop is blocked, the beam that passes through the open interior channel is essentially unaffected by the second stage of the loop: there is nothing left for this beam to recombine with. Hence the device acts merely as a projector for the state labeled by the open channel. Therefore the outcome of the experiment in Figure 6a can be predicted on the basis of the results of Chapter 3. The output beam of A is y'-polarized, and its intensity is I_o (the intensity of the input beam) times the projection probability from state y to state y'. This projection probability has the value $\cos^2 30° = 3/4$. When the beam passes through projector B, its intensity is further fractionally diminished by the projection probability from state y' to state x, that is, $\cos^2 60° = 1/4$. The intensity of the final beam that emerges from B should therefore be $(1/4)(3/4)I_o = (3/16)I_o$. Experiment confirms this prediction.

In the second part of the experiment (Figure 6b), a stop is placed in the y'channel of analyzer loop A instead of the x' channel. This experiment can be analyzed in precisely the same way as the preceding one: the projection probabilities are now 1/4 for the first stage and 3/4 for the second, and the final intensity is again $(3/16)I_o$. This prediction is likewise confirmed by experiment.

In the third part of the experiment (Figure 6c), both channels are opened. In this case, by experiment, no photons whatever are transmitted. In parts one and two we saw that three sixteenths of all incident photons are transmitted by each alternative path alone. Yet when both paths are open no photons at all appear in the output beam. By increasing the number of paths available to photons we have decreased the probability that they are transmitted! How can the concept of photons as particles be consistent with this experimental result?

One simple argument predicts a non-zero output intensity for the recombined beams: A photon in the beam that emerges from B must (so the argument goes) have followed one of two possible paths through the loop, either through the x' channel or through the y' channel. According to the first two parts of the experiment, the probability for a photon to get through by

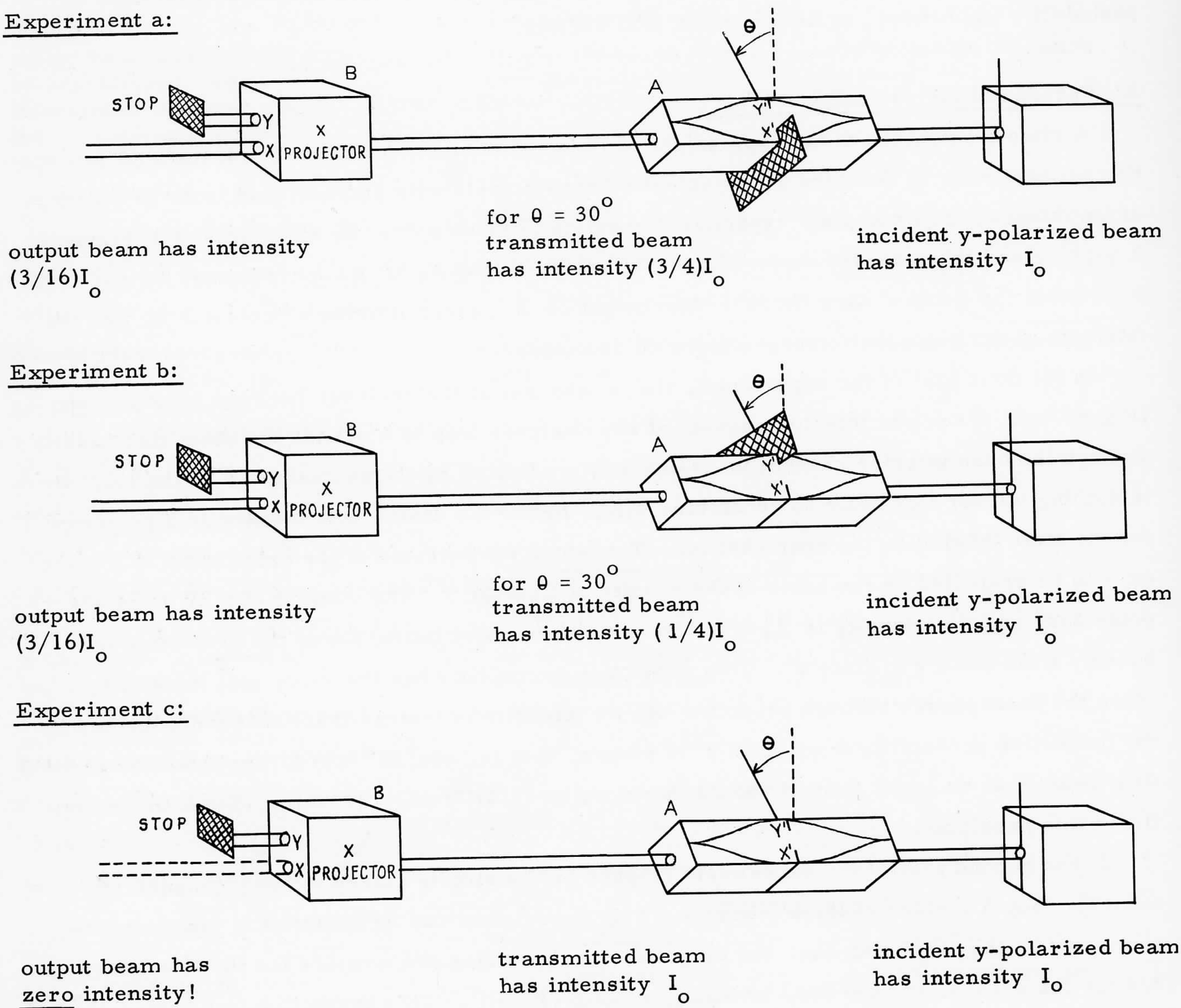


Figure 6. Opening two channels of the analyzer loop rather than one channel results in a decrease in the intensity of the photon beam that emerges from projector B.

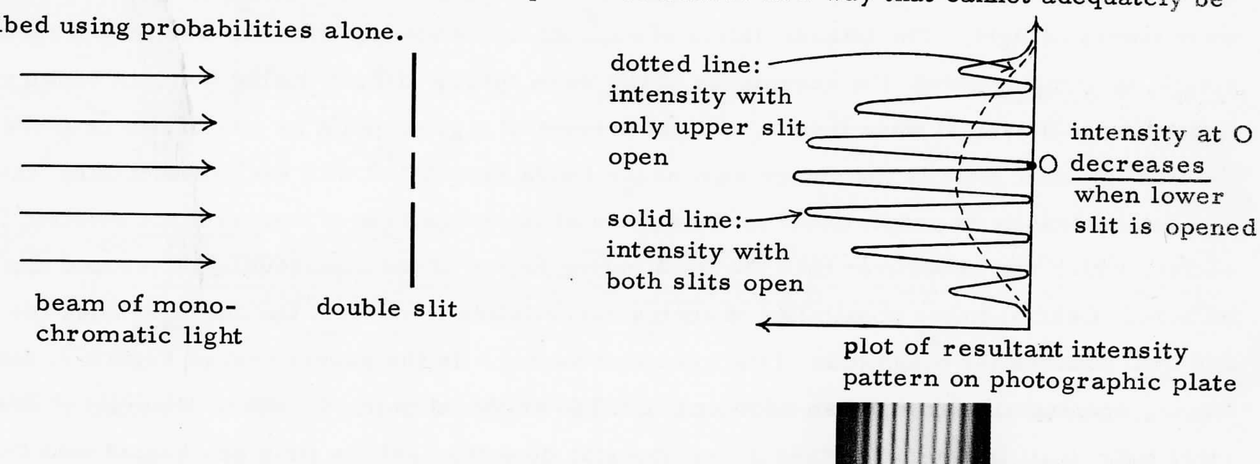
way of channel x' is 3/16; the probability for it to get through by way of channel y' is the same. The total probability for a photon to get through ought to be the sum of the probabilities to get through by the alternate paths, the sum being 3/8. One could conclude, on the basis of this argument, that the intensity of the emerging beam should be $(3/8)I_o$. (If device B is a y projector instead of an x projector, the analysis leads to a prediction of $(5/8)I_o$ as the intensity of the finally-emerging beam.) Yet by experiment the intensity of the emerging beam is zero.

An alternative analysis of the same experiment makes it understandable why no photons emerge in the third step. The analyzer loop is defined by the property that when both channels are open the output beam is identical to the input beam. In Figure 6c the analyzer loop input is y-polarized; therefore the output of the analyzer loop is also y-polarized. The y-polarized beam then falls on an x-projector. Since the projection probability from state y to state x is zero, no photons are transmitted by the final x-projector. (If device B is a y-projector, the entire beam ought to be transmitted.) We therefore arrive at the correct predictions 0 and 1 instead of the incorrect predictions 3/8 and 5/8 given by the earlier line of argument.

The experimental results pose a serious problem for the photon model of light. The incorrect argument above was based on the assumption that the resultant probability with both paths open is the sum of the probabilities for travel by each alternative path. Apparently this assumption must now be reconsidered.

The result we have been describing is one example of a general phenomenon called interference, which occurs when two or more beams derived from the same source are superposed. The "dark" portions of an interference pattern exhibit the same property that has been perplexing us here: they are illuminated less when both beams are superposed than when either beam separately is present. Figure 7 shows a typical interference pattern, produced by light which passes through two narrow slits and impinges on a screen. This experiment was first performed by Thomas Young about 1800.

Interference patterns appear also in properly-designed experiments with particles other than photons. Figure 8 shows an interference pattern obtained with beams of electrons passing through two slits. This experiment is discussed in detail in Chapter 9. Experiments similar in principle to this one have been carried out with beams of other particles, such as neutrons and even helium atoms. On the atomic level all particles behave in a way that cannot adequately be described using probabilities alone.



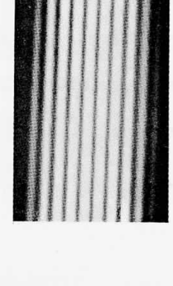


Figure 7. Young's double-slit experiment with light.
Photograph of interference pattern taken by J. W. Cottingham.

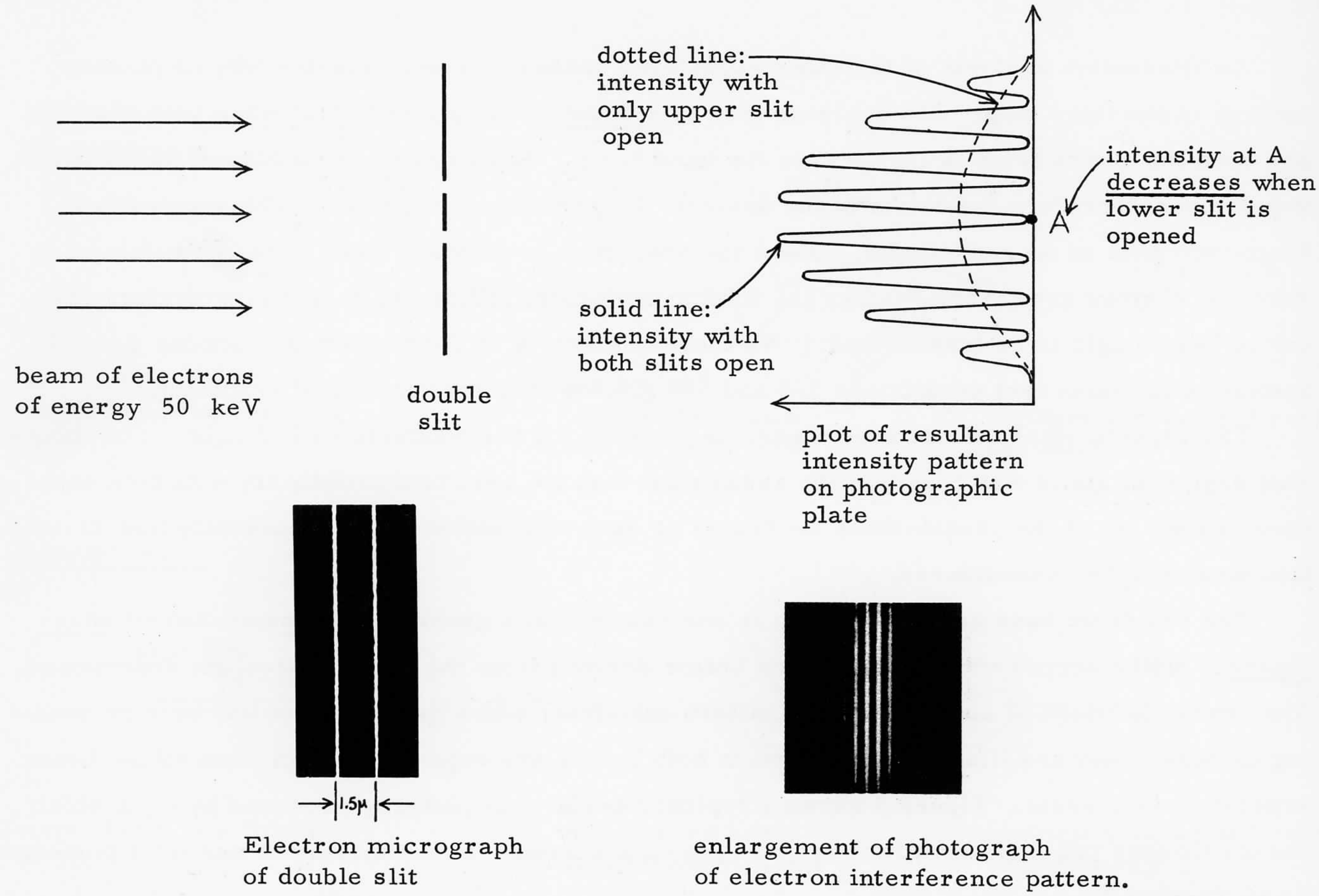
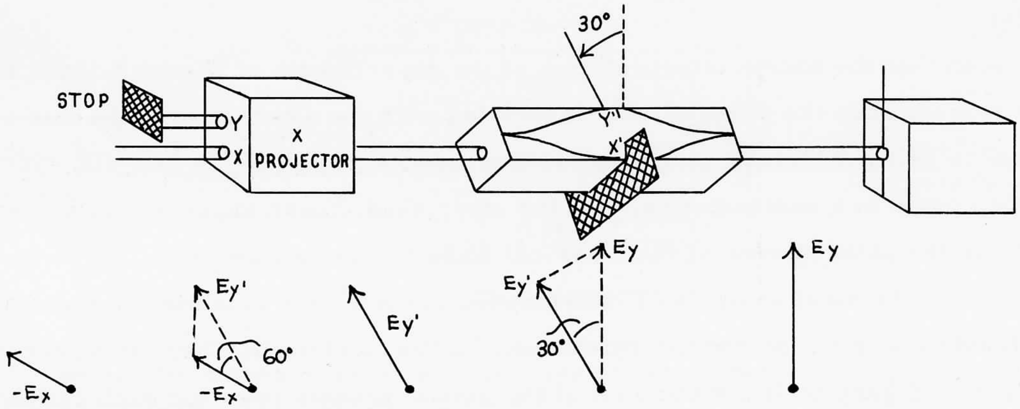
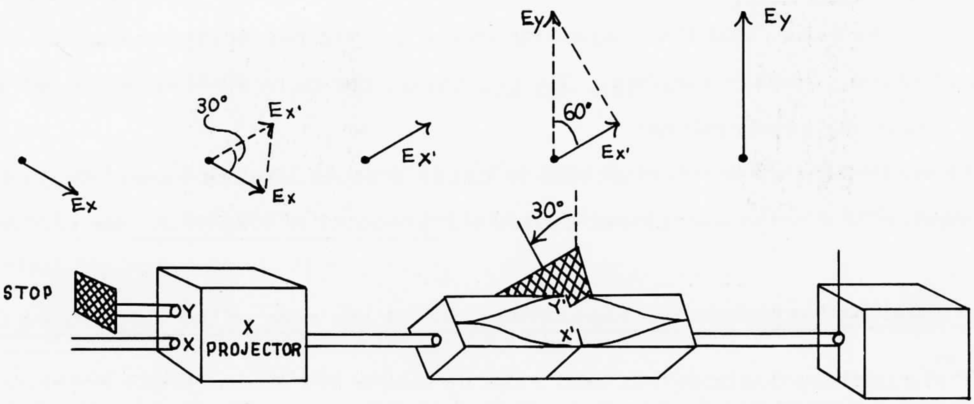


Figure 8. Double slit interference pattern with electrons photographed by Claus Jönssen (Zeitschrift für Physik, 161, 454 (1961)). Schematic diagram above omits numerous electrostatic "lenses" used to enlarge the interference pattern before projection on the photographic plate. The article from which these photographs are taken shows electron interference patterns obtained using 1, 2, 3, 4, and 5 slits.

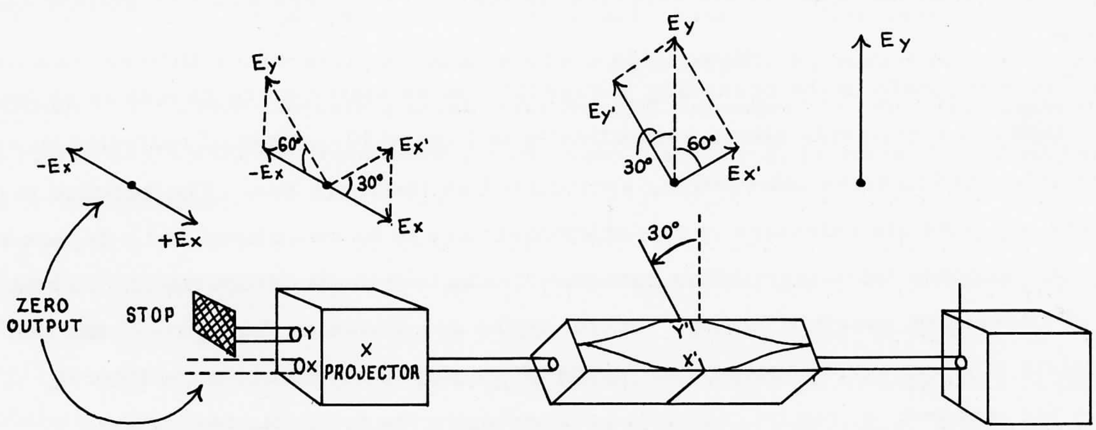
Interference experiments have a straightforward explanation on the basis of the classical wave theory of light. The interpretation of such interference experiments was in fact instrumental in bringing about the acceptance of the wave theory of light during the 19th century. According to classical wave theory, if light arrives at a given point by way of two or more paths, the total electric field is the vector sum of the fields associated with each contributing wave. The total intensity is proportional to the square of the magnitude of the resultant electric field vector, which can be shorter than the field vector of any of the contributing waves and can even be zero. Crucial to the possibility of such a cancellation is the fact the one first adds the vectors and then squares the magnitude of the resultant vector. In the experiment of Figure 7, for example, opening slit 2 causes an additional field to arrive at point O which "interferes destructively with" (that is, points always in the opposite direction as) the field associated with the wave from slit 1. The total electric field therefore vanishes and zero intensity results. A similar argument explains the result of zero intensity in the experiment of Fig. 6c. Figure 9 illustrates the classical wave analysis of this experiment.



Experiment a: y'channel of analyzer loop open



Experiment b: x' channel of analyzer loop open



Experiment c: both channels of analyzer loop open.

Figure 9. Wave optical explanation of the result of experiments of Figure 6. In experiment (c) electric field amplitudes for alternative paths interfere to give zero resultant.

4. Amplitudes

We have seen that the photon interpretation of the experiments of Figure 6 leads to contradiction if one merely adds the probabilities associated with the alternative paths that a photon may travel before being detected. A similar photon interpretation of the two-slit experiment of Figure 7 also leads to a contradiction. On the other hand, these experiments are not at all paradoxical from the point of view of the classical wave theory of light.

The successful classical analysis of interference provides the clue for the resolution of the apparent contradiction in the photon interpretation. If the total probability for a photon to get through the loop in Figure 6c is not the sum of the partial probabilities for each path alone, the two partial probabilities must somehow "interfere" with one another. But probabilities are intrinsically non-negative: the sum of two probabilities can never be less than either one. On the other hand, complex numbers are quantities that share with vectors the mathematical property of interference in the sense that the magnitude of the sum of two complex numbers can be smaller than the magnitude of either number. We can obtain the desired "cancellation" of probability by making the following assumptions:

i) the probability for an incident photon to cause a count in each experiment of Figure 6 is the absolute square of a complex number called the probability amplitude for that measurement and

ii) the probability amplitude for experiment 6c is the sum of the amplitudes for experiments 6a and 6b.

That is, when a photon can follow more than one path, one adds the amplitudes associated with the alternative paths rather than adding the probabilities. This hypothesis accounts for the observed results if the amplitudes for experiments (6a) and (6b) have equal magnitude but opposite sign. *

The hypothesis set forth in the preceding paragraph can be stated more generally as follows. Consider the class of experiments shown symbolically in Figure 10. A set of particles in some initial state is subjected to some interaction, symbolized by the large box. The relative probabilities of various possible outcomes of the experiment are to be calculated and compared with experiment. Suppose that a particular outcome q can be reached from the initial state by a number of alternative possible "paths." (three paths are shown in the figure), and that each path consists of a number of successive "steps" (two steps are shown in the figure). Then the probability for outcome q can be calculated according to the following two rules:

Rule 1. The probability for outcome q is the squared magnitude of a sum of amplitudes for the alternative paths.

$$\begin{pmatrix}\text{probability}\\ \text{for outcome } q\end{pmatrix} = \Big|\sum_i \text{(probability amplitude for path } i)\Big|^2 \qquad (1)$$

* For this particular experiment, it suffices to let the amplitudes be positive or negative real numbers. But, as we shall soon see, in general the amplitudes must be complex.

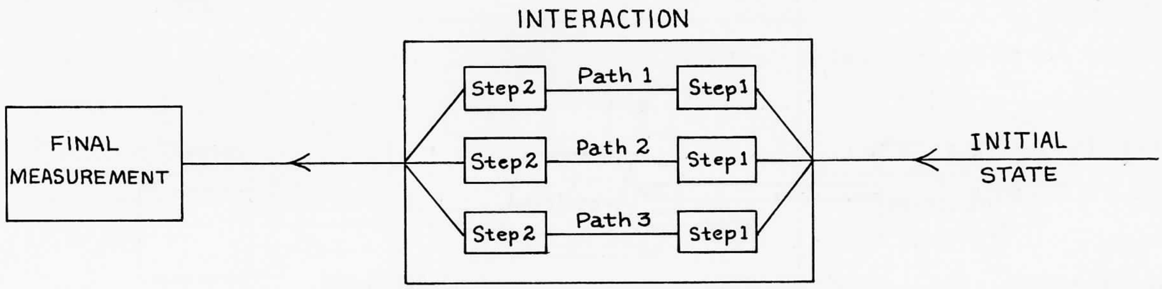


Figure 10. Schematic diagram for a general class of experiments leading from an initial state to a final measurement. The case shown here has three alternative paths from initial state to final measurement and each path contains two steps.

Rule 2. The probability amplitude for each path, say path i, is a product of amplitudes for each step in the path.

$$\begin{pmatrix}\text{probability amplitude} \\ \text{for path } i\end{pmatrix} = \prod_j (\text{amplitude for step } j) \tag{2}$$

The experiment of Figure 6c is clearly of the type symbolized in Figure 10 (with two paths instead of three). In the following section we use probability amplitudes and rules 1 and 2 to analyze in detail the experiment of Figure 6.

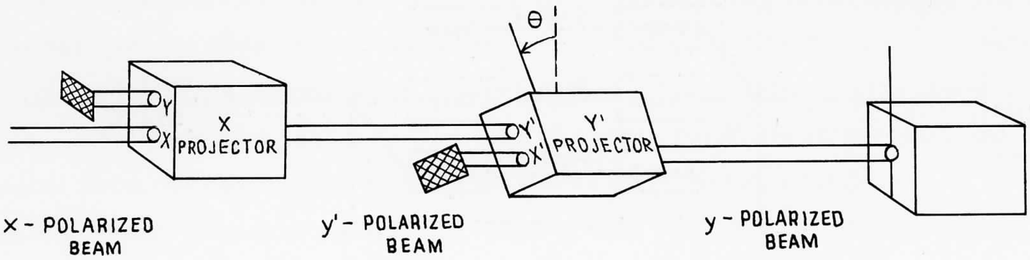
5. Projection amplitudes for states of linear polarization

The only photon experiments for which we have thus far obtained quantitative results are those measuring projection probabilities. Such an experiment is an especially simple special case of the one shown in Figure 10: there is only one path from initial state to final measurement, and only one step in this path. Rules 1 and 2 assert for this case merely that the projection probability is the absolute square of a single amplitude, which we call a projection amplitude. We write projection amplitudes using a convenient bracket notation due to Dirac. The projection amplitude from some initial state, labeled i, to some final state, labeled f, will be written thus:

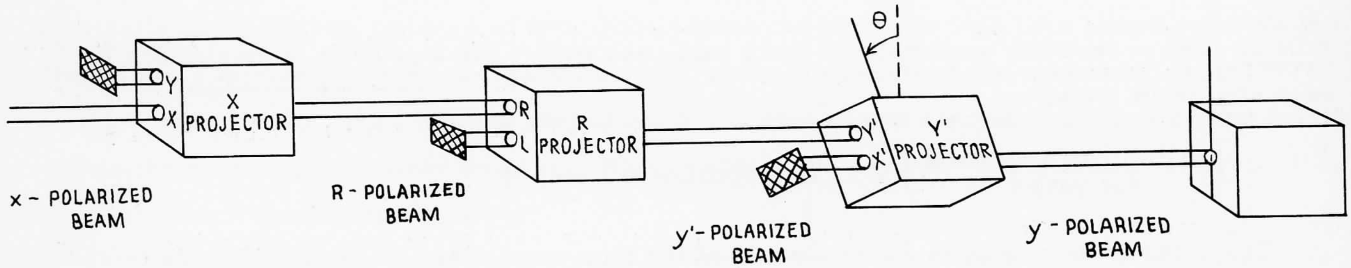
$$\langle f | i \rangle$$

Notice that by convention the initial state is written at the right side of the bracket and the final state at the left.

If an experiment consists of passing a single beam through a succession of projectors, the probability amplitude is, according to rule 2, a product of projection amplitudes. Figure 11 shows a couple of simple examples. Because of the conventional order in which the states are written in a projection amplitude, it is convenient to draw these figures with the incident beam coming from the right. We also write the amplitudes in the same order as the projectors appear in the figure. The absolute square of a product of complex numbers is the product of the absolute squares of the individual numbers. Therefore if the absolute square is taken of the product of amplitudes in the figure, written down according to rule 2, one obtains the usual com-



a) Probability amplitude $= \langle x | y' \rangle \langle y' | y \rangle$



b) Probability amplitude $= \langle x | R \rangle \langle R | y' \rangle \langle y' | y \rangle$

Figure 11. Examples of probabilıty amplitudes for a sequence of projections.

bination law for the resultant probability of sequential events (rule 2 on page 9-7 of Physics, A New Introductory Course, Vol. I).

The magnitudes of the complex numbers $\langle x | x' \rangle$, $\langle y | R \rangle$, etc., are determined from the corresponding projection probabilities, all of which have been measured in Chapter 3; however, we have as yet no information concerning the phases of these numbers. For example, the projection amplitude $\langle x | R \rangle$ could be any number of the form $\frac{1}{\sqrt{2}} e^{i\alpha}$ (as long as α is real), without affecting the outcome of any experiment thus far considered. Furthermore, no sequence of projections, such as the experiments of Figure 11, can give information concerning the phases of projection amplitudes. The product of amplitudes contains an overall phase factor, which disappears when the absolute square is taken (rule 1). Phase information comes only from experiments in which amplitudes interfere, i.e., add. The experiment which led us to introduce the idea of amplitudes, that of Figure 6c, is just such a one. By applying rules (1) and (2) to the analysis of this experiment, as well as other similar ones, we shall infer some information concerning the relative phases of various projection amplitudes. We shall not be able to derive a unique value for each phase. Nevertheless, the information we obtain on relative phases turns out to be sufficient to predict correctly the outcome of all other pertinent experiments.

Now we derive expressions for the projection amplitudes by applying rules 1 and

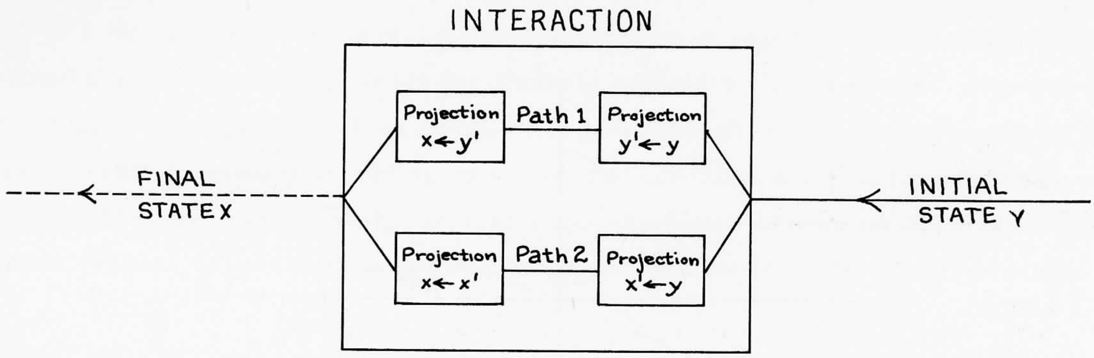


Figure 12. Schematic diagram of the experiment of Figure 6c. Compare this diagram with Figure 10.

2 to the experiment of Figure 6c. There are two paths in the experiment, and two steps in each path. Figure 12 shows the experiment symbolized in the manner of Figure 10. Each step may be considered a projection, so we can write the probability amplitude for the experiment as

$$\text{probability amplitude} = \langle x|y'\rangle\langle y'|y\rangle + \langle x|x'\rangle\langle x'|y\rangle \qquad (3)$$

The experimental result (zero counting rate with both channels open) indicates that the absolute square of the probability amplitude (3) vanishes, and therefore that the amplitude itself vanishes:

$$\langle x|y'\rangle\langle y'|y\rangle + \langle x|x'\rangle\langle x'|y\rangle = 0 \qquad (4)$$

The square of the magnitude of each projection amplitude that appears in Eq. 4 is the projection probability between the designated states. According to the results of Chapter 3, each projection probability is the squared cosine of the angle between the axes that define the two states involved. We can express all these in terms of θ, the angle between the y and y' axes. (See Figure 13.)

$$|\langle x|x'\rangle|^2 = |\langle y'|y\rangle|^2 = \cos^2\theta \qquad (5)$$

$$|\langle x'|y\rangle|^2 = |\langle x|y'\rangle|^2 = \sin^2\theta \qquad (6)$$

The general solutions of equations (5) and (6) can be written as real magnitudes times complex phase factors of the form $e^{i\phi}$, whose magnitude is unity. (It is conventional to call ϕ a phase, and $e^{i\phi}$ a phase factor.)

$$\begin{aligned}
\langle y'|y\rangle &= \cos\theta\ e^{i\alpha} \\
\langle x|x'\rangle &= \cos\theta\ e^{i\beta} \\
\langle x'|y\rangle &= \sin\theta\ e^{i\gamma} \\
\langle x|y'\rangle &= \sin\theta\ e^{i\delta}
\end{aligned} \qquad (7)$$

where α, β, γ, and δ are real. We are interested in the values of the phase factors $e^{i\alpha}$, $e^{i\beta}$, $e^{i\gamma}$, and $e^{i\delta}$. (These values could depend on θ.) Substituting the forms (7) into (4) gives

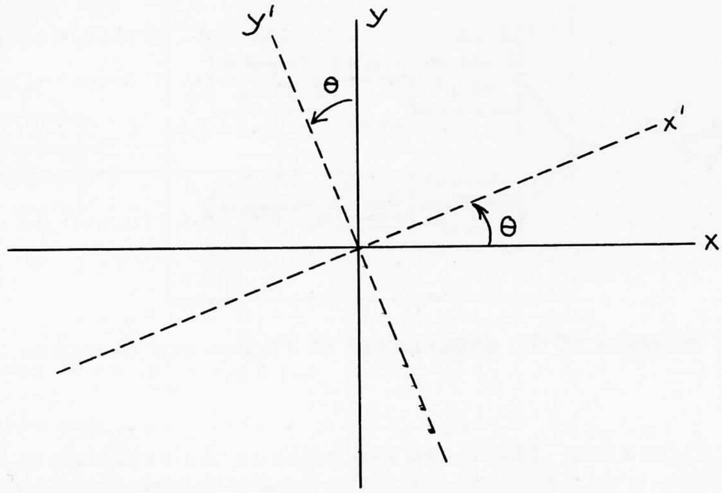


Figure 13. Angles used in projection probabilities.

$$\cos\theta \sin\theta (e^{i\beta} e^{i\gamma} + e^{i\alpha} e^{i\delta}) = 0 \tag{8}$$

which implies, as long as $\cos\theta \sin\theta \neq 0$, that

$$e^{i(\beta+\gamma)} = -e^{i(\alpha+\delta)} \tag{9}$$

The experiment under study provides only one relation, Eq. 9, among the four phases; any set of phases that satisfies Eq. 9 leads to a correct prediction of the result of this experiment. Consequently we are very far from being able to determine each of the phase factors uniquely. Other experiments, in which the same projection amplitudes appear, might lead to further restrictions on the phases, or perhaps even determine them unambiguously. But no such experiment has ever been provided. Apparently the results of all experiments with linearly polarized photons are independent of the choice of phases for the amplitudes (7), provided that Eq. 9 is satisfied. This is an example of a general result in the study of quantum amplitudes: only certain limited information concerning phases may be obtained from experiment. Moreover, what limited information _is_ available from experiment always concerns relative phases between amplitudes rather than the absolute phase of any one amplitude. But this information is always sufficient for predicting the results of other experiments in which the same amplitudes appear.

For the present problem it is possible to choose a particularly simply phase convention, namely one in which all amplitudes are real and the phase factors in (7) are independent of θ. Notice that Eq. 9 is satisfied if any three of the four phase factors $e^{i\alpha}$, $e^{i\beta}$, $e^{i\gamma}$, $e^{i\delta}$ are equal to +1 and the fourth is equal to -1. (or, alternatively, if any three are -1 and the fourth is +1.) It is immaterial which phase factor is chosen to be the negative one. We adopt the following choice of phases:

$$\langle x | x' \rangle = \langle y' | y \rangle = \cos\theta \tag{10}$$

$$\langle x' | y \rangle = \sin\theta \tag{11}$$

$$\langle x | y' \rangle = -\sin\theta \tag{12}$$

There is a simple way to summarize the combination of experiment and convention by which we have fixed the projection amplitudes for linearly polarized photon states. Experiment determines the polarization axis of a given beam, with no unique direction either way along this axis. The phase convention amounts to adding an arrow that specifies a particular direction along this axis. We can if we like permanently inscribe such coordinate arrows on the analyzer that produces a state (Figure 14). Then the projection amplitudes for linearly polarized states (accord-

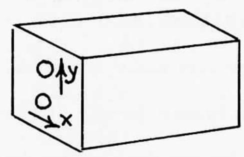


Figure 14. Analyzer with inscribed coordinate arrows.

ing to the convention we have adopted) can be summarized by one simple rule. For states defined by particular channels of particularly oriented analyzers, the projection amplitude is just the cosine of the angle between corresponding arrows. One example is shown in Figure 15. A positive angle is defined by a counterclockwise rotation from the initial to the final state in the projection. It may be verified directly that the amplitudes given by Eqs. 10 to 12 are all consistent with this simple rule--this is, of course, the reason we chose the phase convention as we did. The rule also implies that for our phase convention the following relation holds for any two states of linear polarization j and k:

$$\langle j | k \rangle = \langle k | j \rangle \quad \text{(linear polarization states)} \qquad (13)$$

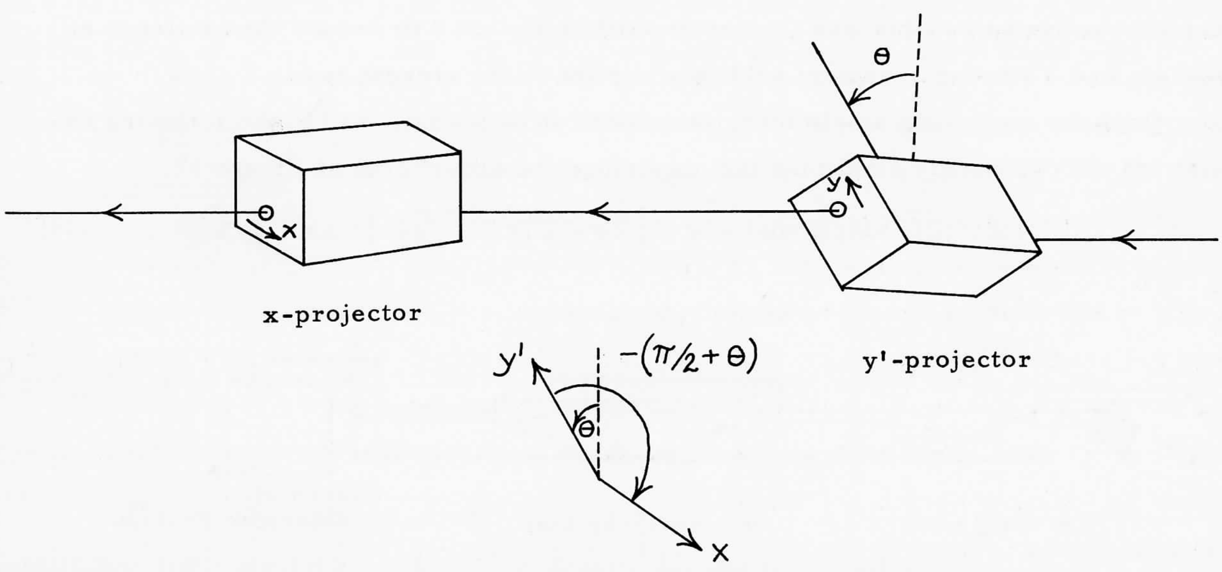


Figure 15. Projection amplitudes for photon states of linear polarization can be simply calculated as the cosine of the angle between coordinate arrows of the corresponding projectors. This procedure automatically chooses the phase convention adopted in the text. For the example shown (see Eq. 12) $\cos -(\pi/2 + \theta) = -\sin\theta$

This equation results from the fact that interchanging initial and final states is equivalent to changing the sign of the angle between them, and the cosine is an even function. The states on the left and right sides of (13) must of course be defined by identically oriented analyzers.

6. Projection amplitudes involving states of circular polarization

In this section we investigate the amplitudes for projection from states of linear polarization to states of circular polarization, and vice versa. The need for such amplitudes can be demonstrated by considering experiments similar to those of Figure 6 (Sec. 3), with an RL analyzer loop replacing the x'y' analyzer loop. (The RL analyzer loop is constructed from two RL analyzers in the same way as the xy analyzer loop is constructed from two linear analyzers. An RL analyzer loop is shown schematically in Figure 16.) Just as in the first set of

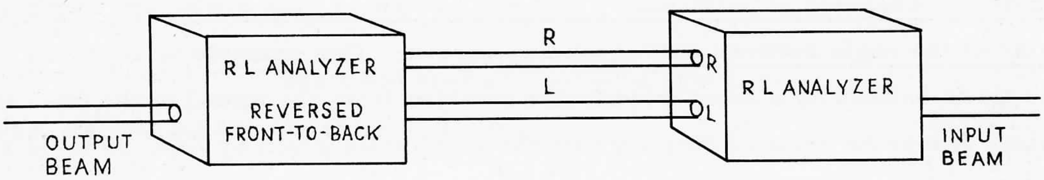


Figure 16. An RL analyzer loop.

experiments, a beam emerges from the final projector when either channel of the analyzer loop is blocked, but no photons emerge when both channels are open. It follows that the probability of detecting a photon when both channels are open cannot be the sum of the probabilities associated with the alternative paths; this was the result used in Section 4 to deduce the existence of amplitudes, and a similar argument evidently applies to the present case.

The rules for combining amplitudes, enunciated in Section 4, lead to the following expression for the probability amplitude that describes the experiment of Figure 17

$$(\text{probability amplitude}) = \langle y' | R \rangle \langle R | y \rangle + \langle y' | L \rangle \langle L | y \rangle \qquad (14)$$

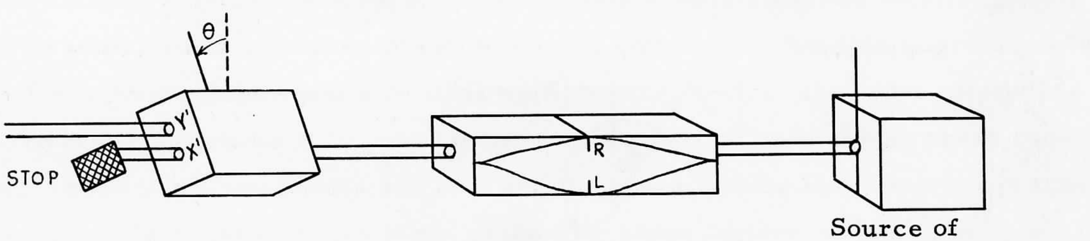


Figure 17. Experiment to determine projection amplitudes for circularly polarized photon beams.

Since by definition the analyzer loop with both channels open has no effect on the beam incident on it, the probability amplitude for this experiment must be $\langle y' | y \rangle$. According to the phase convention adopted in the preceding section, this amplitude has the value cos θ.

$$\langle y'|R\rangle\langle R|y\rangle + \langle y'|L\rangle\langle L|y\rangle = \cos\theta \qquad (15)$$

The absolute squares of the four projection amplitudes that appear in Eq. (15) are the corresponding projection probabilities. In Chapter 3, all four of these projection probabilities were found to have the value 1/2. Hence we can write (cf. Eq. 7)

$$\begin{aligned} \langle y'|R\rangle &= e^{i\alpha}/\sqrt{2} \\ \langle R|y\rangle &= e^{i\beta}/\sqrt{2} \\ \langle y'|L\rangle &= e^{i\gamma}/\sqrt{2} \\ \langle L|y\rangle &= e^{i\delta}/\sqrt{2} \end{aligned} \qquad (16)$$

Substituting these forms in Eq. (15), we obtain:

$$\frac{1}{2}(e^{i\alpha}e^{i\beta} + e^{i\gamma}e^{i\delta}) = \cos\theta = \frac{1}{2}(e^{i\theta} + e^{-i\theta}) \qquad (17)$$

Eq. 17 is the only relation that can be found among the phase factors. Just as in the analogous situation discussed in the preceding section, there are many possible sets of phases that satisfy (17). However, in the present case it is not possible to find a solution in which all the phase factors are real, (Except when θ is 0 or $\pi/2$): the amplitudes that relate linear and circular polarization states are necessarily complex. The following phase convention turns out to be most convenient.

$$\langle y'|R\rangle = e^{i\theta}/\sqrt{2} \qquad (18a)$$

$$\langle R|y\rangle = 1/\sqrt{2} \qquad (18b)$$

$$\langle y'|L\rangle = e^{-i\theta}/\sqrt{2} \qquad (18c)$$

$$\langle L|y\rangle = 1/\sqrt{2} \qquad (18d)$$

Once these amplitudes have been fixed, all others are uniquely determined. For example, to find the values of $\langle R|y'\rangle$ and $\langle L|y'\rangle$, consider the experiment of Figure 18:

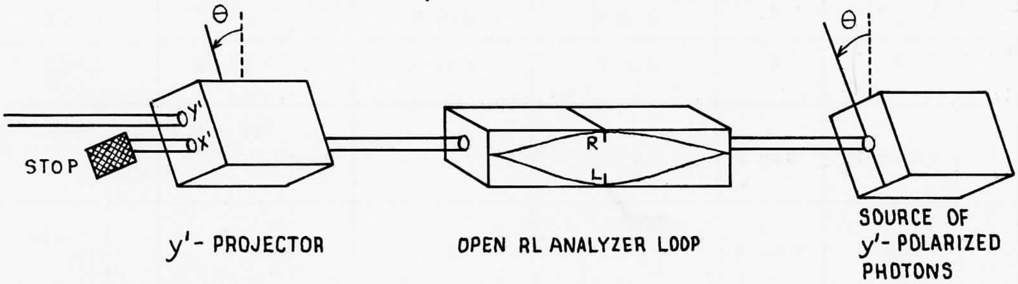


Figure 18. Experiment used to derive the probability amplitudes $\langle R|y'\rangle$ and $\langle L|y'\rangle$

The probability amplitude that describes the experiment has the form

$$(\text{probability amplitude}) = \langle y'|R\rangle\langle R|y'\rangle + \langle y'|L\rangle\langle L|y'\rangle = \langle y'|y'\rangle = 1 \qquad (19)$$

Substituting for $\langle y'|R\rangle$ and $\langle y'|L\rangle$ from (18a) and (18c), we get

$$\frac{1}{\sqrt{2}}(e^{i\theta}\langle R|y'\rangle + e^{-i\theta}\langle L|y'\rangle) = 1 \qquad (20)$$

Moreover, $\langle R|y'\rangle$ and $\langle L|y'\rangle$ must each have magnitude $1/\sqrt{2}$. The only solution of (20) consistent with this condition is

$$\langle R | y' \rangle = e^{-i\theta}/\sqrt{2} \tag{21a}$$

$$\langle L | y' \rangle = e^{+i\theta}/\sqrt{2} \tag{21b}$$

Equations (18a, c) and (21 a, b) actually specify all possible amplitudes between linear and circular states inasmuch as y' denotes an arbitrary state of linear polarization. In particular, putting $\theta = 0$ in (21 a, b) reproduces (18 b, d); putting $\theta = 0$ in (18 a, c) gives

$$\langle y | R \rangle = 1/\sqrt{2} \tag{22a}$$

$$\langle y | L \rangle = 1/\sqrt{2} \tag{22b}$$

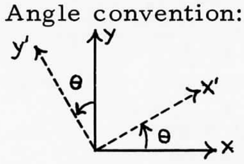
Finally, we can let $\theta + 3\pi/2$, which amounts to putting y' along the x axis. This gives

$$\begin{aligned} \langle x | R \rangle &= -i/\sqrt{2} \\ \langle x | L \rangle &= i/\sqrt{2} \\ \langle R | x \rangle &= i/\sqrt{2} \\ \langle L | x \rangle &= -i/\sqrt{2} \end{aligned} \tag{23}$$

The fact that the amplitudes $\langle y | L \rangle$, $\langle R | y \rangle$, etc. are all real while $\langle x | L \rangle$, $\langle R | x \rangle$, etc. are pure imaginary is of course a consequence of the particular phase convention we have adopted, and carries no further significance. A summary of the probability amplitudes for photon polarization states derived thus far is presented in Table 1.

Table 1: Projection amplitudes for photon polarization states

Angle convention: (Beam emerges toward reader.)



FROM STATE

TO STATE	$\|x\rangle$	$\|y\rangle$	$\|x'\rangle$	$\|y'\rangle$	$\|R\rangle$	$\|L\rangle$
$\langle x\|$	1	0	$\cos\theta$	$-\sin\theta$	$-i/\sqrt{2}$	$i/\sqrt{2}$
$\langle y\|$	0	1	$\sin\theta$	$\cos\theta$	$1/\sqrt{2}$	$1/\sqrt{2}$
$\langle x'\|$	$\cos\theta$	$\sin\theta$	1	0	$\frac{1}{\sqrt{2}}e^{i(\theta - \frac{\pi}{2})}$	$\frac{1}{\sqrt{2}}e^{-i(\theta - \frac{\pi}{2})}$
$\langle y'\|$	$-\sin\theta$	$\cos\theta$	0	1	$\frac{1}{\sqrt{2}}e^{i\theta}$	$\frac{1}{\sqrt{2}}e^{-i\theta}$
$\langle R\|$	$i/\sqrt{2}$	$1/\sqrt{2}$	$\frac{1}{\sqrt{2}}e^{-i(\theta-\frac{\pi}{2})}$	$\frac{1}{\sqrt{2}}e^{-i\theta}$	1	0
$\langle L\|$	$-i/\sqrt{2}$	$1/\sqrt{2}$	$\frac{1}{\sqrt{2}}e^{i(\theta - \frac{\pi}{2})}$	$\frac{1}{\sqrt{2}}e^{i\theta}$	0	1

Notice that with our phase conventions, all amplitudes satisfy the condition

$$\langle j | k \rangle = \langle k | j \rangle^* \tag{24}$$

where j and k are arbitrary states: interchanging initial and final states turns each amplitude into its complex conjugate. This property is quite generally true of quantum amplitudes. That is, a phase convention can always be picked so that (24) is satisfied. The projection probabilities are the same when initial states are exchanged.

$$|\langle j|k\rangle|^2 = |\langle k|j\rangle^*|^2 = |\langle k|j\rangle|^2$$

For states of linear polarization we found that probability amplitudes remain the same when initial and final states are exchanged. This is a special case of Eq. 24, inasmuch as the amplitudes between states of linear polarization are all real according to the conventions we have adopted.