CRC Press
Taylor & Francis Group

LECTURES ON QUANTUM MECHANICS

Gordon Baym

LECTURE NOTES AND SUPPLEMENTS IN PHYSICS

ABP

LECTURES ON QUANTUM MECHANICS

LECTURE NOTES AND SUPPLEMENTS IN PHYSICS

John David Jackson and David Pines, *Editors*

Gabriel Barton	Introduction to Dispersion Techniques in Field Theory, 1965
Gordon Baym	Lectures on Quantum Mechanics, 1969 (3rd printing, with corrections, 1974)
Hans A. Bethe and Roman W. Jackiw	Intermediate Quantum Mechanics, 1968 (2nd printing, with corrections, 1973)
David Bohm	The Special Theory of Relativity, 1965
B. H. Bransden	Atomic Collision Theory, 1970
Willem Brouwer	Matrix Methods in Optical Instrument Design, 1964
R. Hagedorn	Relativistic Kinematics: A Guide to the Kinematic Problems of High-Energy Physics, 1964 (3rd printing, with corrections, Spring 1973)
John David Jackson	Mathematics for Quantum Mechanics: An Introductory Survey of Operators, Eigenvalues, and Linear Vector Spaces, 1962
Robert S. Knox and Albert Gold	Symmetry in the Solid State, 1964
K. Nishijima	Fields and Particles: Field Theory and Dispersion Relations, 1969
David Park	Introduction to Strong Interactions: A Lecture-Note Volume, 1966
David Pines	Elementary Excitations in Solids: Lectures on Phonons, Electrons, and Plasmons, 1964
R. D. Sard	Relativistic Mechanics: Special Relativity and Classical Particle Dynamics, 1970

LECTURES ON QUANTUM MECHANICS

GORDON BAYM

University of Illinois

ABP *Advanced Book Program*

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

First published 1969 by Westview Press

Published 2018 by CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

CRC Press is an imprint of the Taylor & Francis Group, an informa business

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Library of Congress Catalog Card Number: 68–5611

ISBN 13: 978-0-8053-0667-5 (pbk)

PREFACE

These lecture notes comprise a three-semester graduate course in quantum mechanics given at the University of Illinois. There are a number of texts which present the basic topics very well; but since a fair quantity of the material discussed in my course was not available to the students in elementary quantum mechanics books, I was asked to prepare written notes. In retrospect these lecture notes seemed sufficiently interesting to warrant their publication in this format. The notes, presented here in slightly revised form, constitute a self-contained course in quantum mechanics from first principles to elementary relativistic one-particle mechanics. The student may want to look as well at one or more of the standard texts, such as K. Gottfried, *Quantum Mechanics* (W.A. Benjamin, New York, 1966); A. Messiah, *Quantum Mechanics* (North-Holland Publishing Company, Amsterdam, 1961), in two volumes; E. Merzbacher, *Quantum Mechanics* (John Wiley and Sons, New York, 1961); and for relativistic quantum mechanics, J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill Book Company, New York, 1964) and J. Sakurai, *Advanced Quantum Mechanics* (Addison-Wesley Publishing Company, Reading, Massachusetts, 1967). I occasionally refer in the notes to these books by author only. References to the Feynman Lectures are to R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley Publishing Company, Reading, Massachusetts, 1965), in three volumes.

Prerequisite to reading these notes is some familiarity with elementary quantum mechanics, at least at the undergraduate level; preferably the reader should already have met the uncertainty principle and the concept of a wave function. Mathematical prerequisites include sufficient acquaintance with complex variables to be able to do simple contour integrals and to understand words such as "poles" and "branch cuts." An elementary knowledge of Fourier transforms and series is necessary. I also assume an awareness of classical electrodynamics on the level of J. D. Jackson, *Classical Electrodynamics* (John Wiley and Sons, New York, 1962). Finally, I should mention that the figures in the notes are all sketches, not accurate graphs.

Among past and present colleagues I am particularly indebted, for their

patient criticisms, discussions, and comments, to R.W. Hellwarth, L.P. Kadanoff, and D.G. Ravenhall; and to J.D. Jackson, who read the notes through and made many helpful suggestions about revising for publication. I am grateful to Mrs. Marie Brosig who typed the original notes, to W. F. Saam for proofreading this volume, and to Nina Baym, for substantial editorial assistance.

GORDON BAYM

Urbana, Illinois
January, 1968

CONTENTS

Chapter 1
PHOTON POLARIZATION

In order to become more familiar with the concepts and techniques of quantum mechanics, let us concentrate on one of the simplest quantum mechanical systems — one with which we are all very familiar — the polarization of light.

A classical light wave propagating in the z direction is described by the electric field vector

$$\mathbf{E}(r,t) = \begin{pmatrix} E_x(r,t) \\ E_y(r,t) \\ 0 \end{pmatrix}, \tag{1-1}$$

since the electric (and magnetic) field vector is perpendicular to the direction of propagation of the light. In Gaussian units, $|\mathbf{H}| = |\mathbf{E}|$, $H_x = -E_y$, and $H_y = E_x$:

$$\mathbf{H} = \hat{\mathbf{z}} \times \mathbf{E}$$

Thus knowing E we know H.

Since electric fields are real we may write

$$E_x(r,t) = E_x^0 \cos(kz - \omega t + \alpha_x)$$

$$E_y(r,t) = E_y^0 \cos(kz - \omega t + \alpha_y), \tag{1-2}$$

where k is the wavenumber of the light,

$$k = \frac{2\pi}{\lambda}, \tag{1-3}$$

and ω is the angular frequency. α_x and α_y are the phases and E_x^0 and E_y^0 the (real) amplitudes of the electric field components.

It is often more convenient to use a complex notation for the electric field. We define the complex $\mathbf{E}$ vector by $\mathbf{E} = (E_x, E_y, 0)$ where

$$E_x = E_x^0 e^{i\alpha x}, \qquad E_y = E_y^0 e^{i\alpha y}, \tag{1-4}$$

and write

$$E_x(\mathbf{r}, t) = E_x e^{ikz - i\omega t}, \qquad E_y(\mathbf{r}, t) = E_y e^{ikz - i\omega t}, \tag{1-5}$$

remembering that in the end we should take the real part of the complex field.

The polarization state of the light is directly related to the $\mathbf{E}$ vectors. For example:

(i) if $E_y = 0$, the wave is plane polarized in the x direction.

(ii) if $E_x = 0$, the wave is plane polarized in the y direction.

(iii) if $E_x = E_y$, the wave is polarized at 45°, as in Fig. 1-1;

(iv) if $E_y = e^{i\pi/2} E_x = iE_x$, then the y component lags the x component by 90° and the wave is right circularly polarized:

$$\mathrm{Re}\, E_y(rt) \sim \cos(kz - \omega t + \frac{\pi}{2}), \qquad \mathrm{Re}\, E_x(rt) \sim \cos(kz - \omega t).$$

(v) similarly, if $E_y = -iE_x$, the wave is left circularly polarized.

Let us calculate the energy of the wave in terms of $\mathbf{E}$. The energy density of an electromagnetic field is given, in Gaussian units, by

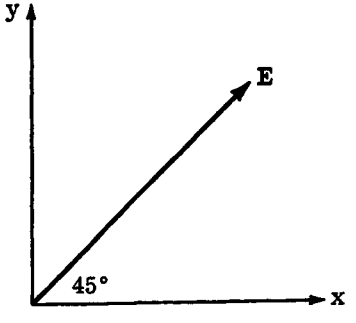

Fig. 1-1

The electric field for 45° polarization.

$$\varepsilon(rt) = \frac{1}{8\pi}[(\mathrm{Re}\,E_x(rt))^2 + (\mathrm{Re}\,E_y(rt))^2 + (\mathrm{Re}\,H_x(rt))^2 + (\mathrm{Re}\,H_y(rt))^2]. \tag{1-6}$$

But for a plane light wave:

$$\mathrm{Re}\,H_x(rt) = -\mathrm{Re}\,E_y(rt), \qquad \mathrm{Re}\,H_y(rt) = \mathrm{Re}\,E_x(rt). \tag{1-7}$$

Thus

$$\varepsilon(rt) = \frac{1}{4\pi}[(\mathrm{Re}\,E_x(rt))^2 + (\mathrm{Re}\,E_y(rt))^2]$$

$$= \frac{1}{4\pi}[|E_x|^2 \cos^2(kz - \omega t + \alpha_x) + |E_y|^2 \cos^2(kz - \omega t + \alpha_y)]. \tag{1-8}$$

Now let us assume that the wave occupies a volume V which is very many wavelengths long in the z direction. The total energy of the wave is thus

$$\varepsilon_{\mathrm{total}} = \int_V d^3r\, \varepsilon(rt) = \frac{1}{4\pi} \cdot \frac{V}{2}(|E_x|^2 + |E_y|^2) = \frac{|\mathbf{E}|^2}{8\pi} V, \tag{1-9}$$

and $|\mathbf{E}|^2/8\pi$ is the average energy per unit volume of the wave.

Consider what happens if we take a wave polarized at a 45° angle to the x axis and pass it through a *polaroid filter* that passes x polarized light, but not y polarized light. Then before the light passes through the polaroid

$$E_x = E_y = E \tag{1-10}$$

and after it passes through,

$$E_x = E, \qquad E_y = 0; \tag{1-11}$$

the beam comes out polarized in the x direction and its total energy is halved. The emerging electric field must be along $\hat{\mathbf{x}}$ since we know that a second such polaroid, whose axis is parallel to the first, would have no further effect on the beam.

Now let us reconsider the effect of the polaroid from a quantum mechanical point of view. First of all we know that in quantum

mechanics the total energy of a wave of frequency ω cannot be arbitrary but must be an integral multiple of $\hbar\omega$:

$$\varepsilon_{\text{total}} = N\hbar\omega \tag{1-12}$$

where N is the number of photons in the wave. Thus when the energy of the wave is halved by the polaroid what must happen is that half the photons pass through and half don't. This is very weird — classically we would say that if any photon gets through, then since all photons are identical and all see identical conditions at the polaroid they should all pass through. The only way we can interpret what happens at the polaroid is to say that each photon has a *probability* one-half of passing through. We are really forced into a probabilistic point of view by the fact that the energy of electromagnetic radiation is quantized!

An immediate consequence of the fact that the passage of photons through the polaroid is governed by the laws of probability is that only rarely will *exactly* half the photons pass through. The mean number of photons passing through will be half the incident number — this must be the case if in the classical limit, i.e., when the beam consists of many, many photons, we are to recover from the probability laws the "deterministic" classical law for the passage of light through the polaroid. However, there will always be fluctuations about this mean number due to the finite size of $\hbar$. (The requirement that quantum mechanics yields the correct classical limit is called the *correspondence principle.)*

In general, to calculate the probability of the photon passing through the polaroid, we just have to ask for the fraction of the energy of a similar classical beam that is passed by the polaroid. This fraction is given by the ratio

$$\frac{|E_x|^2}{|E_x|^2 + |E_y|^2} = \frac{|E_x|^2}{|\mathbf{E}|^2} \tag{1-13}$$

For example, if the beam is polarized at an angle θ to the x axis, then $|E_x| = |\mathbf{E}| \cos\theta$. Hence a fraction $\cos^2\theta$ of the total energy passes through, and we would conclude that a single photon whose polarization vector was at an angle θ to the x axis would have a probability $\cos^2\theta$ of passing through an x-polaroid. Note that the photon, *if* it passes through, emerges, according to (1-11) polarized in the x direction!

Similarly, if we had a prism that passed only right circularly polarized light, then to calculate the fraction of the energy passed, we would write the beam as a coherent superposition of right

circularly polarized light plus left circularly polarized light

$$\mathbf{E} = \mathbf{E}_{RCP} + \mathbf{E}_{LCP}; \tag{1-14}$$

Then the effect of the prism would be to throw away the $\mathbf{E}_{LCP}$ component and pass $\mathbf{E}_{RCP}$. The fraction of energy passed is thus

$$\frac{|\mathbf{E}_{RCP}|^2}{|\mathbf{E}_{RCP}|^2 + |\mathbf{E}_{LCP}|^2}, \tag{1-15}$$

and quantum mechanically we would interpret this fraction as the probability that one photon would pass through the prism, and emerge with right circular polarization.

Since all beams of light are superpositions of many beams consisting of one photon each, we shall turn our attention to the polarization properties of single photons. As we have already seen, it will be easy to discover the probability rules for one photon from our knowledge of the behavior of classical beams. The general laws of quantum mechanics are just generalizations of these rules.

For one photon, we have, from (1-9) and (1-12), that

$$|\mathbf{E}|^2 V = 8\pi\hbar\omega. \tag{1-16}$$

We shall define the *state vector* of the photon polarization

$$|\Psi\rangle = \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix}, \tag{1-17}$$

by writing

$$\psi_x = \sqrt{\frac{V}{8\pi\hbar\omega}}\, E_x, \qquad \psi_y = \sqrt{\frac{V}{8\pi\hbar\omega}}\, E_y. \tag{1-18}$$

The $|\Psi\rangle$ vectors are vectors in a *complex* two-dimensional space, since their components are complex numbers. From (1-16) it follows at once that $|\Psi\rangle$ has unit length:

$$|\psi_x|^2 + |\psi_y|^2 = 1. \tag{1-19}$$

In fact, the state vectors are independent of the volume V and depend only on the state of polarization of the photon. For example, if

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{i\alpha} \\ e^{i\alpha} \end{pmatrix},$$

then the photon is polarized at 45° to the x axis. A knowledge of the $|\Psi\rangle$ vector gives us all the information we can have about the state of polarization of the photon.

Some special examples of these vectors are

$|x\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$: x polarization

$|y\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$: y polarization

$|R\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$: right circular polarization

$|L\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}$: left circular polarization.

Let us associate with each column vector $|\Psi\rangle$ a row vector $\langle\Psi|$ which we define by

$$\langle\Psi| = (\psi_x^* \quad \psi_y^*), \tag{1-20}$$

where * stands for complex conjugate. Also we shall define the *scalar product* of a row vector $\langle\Phi|$ and a column vector $|\Psi\rangle$ to be

$$\langle\Phi|\Psi\rangle = \phi_x^*\psi_x + \phi_y^*\psi_y = \langle\Psi|\Phi\rangle^*. \tag{1-21}$$

The normalization condition (1-19) on the $|\Psi\rangle$ vectors can thus be written as

$$\langle\Psi|\Psi\rangle = 1. \tag{1-22}$$

Clearly

$$\langle x|x\rangle = 1 = \langle y|y\rangle, \qquad \langle R|R\rangle = 1 = \langle L|L\rangle. \tag{1-23}$$

The vectors $|x\rangle$ and $|y\rangle$ are *orthogonal,* i.e., perpendicular, in the sense that

$$\langle x|y\rangle = 0. \tag{1-24}$$

They are said to form a *basis,* since any $|\Psi\rangle$ vector can be written as a linear superposition of them

$$|\Psi\rangle = \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} = \psi_x|x\rangle + \psi_y|y\rangle. \tag{1-25}$$

Because they are orthogonal and satisfy the normalization condition (1-22), the basis they form is called *orthonormal.* Similarly, the set $|R\rangle$ and $|L\rangle$ form an orthonormal basis, since we can always write

$$|\Psi\rangle = \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} = \frac{\psi_x - i\psi_y}{\sqrt{2}}|R\rangle + \frac{\psi_x + i\psi_y}{\sqrt{2}}|L\rangle. \tag{1-26}$$

If we take the scalar product of both sides of Eq. (1-25) with $\langle x|$ we see that

$$\langle x|\Psi\rangle = \psi_x\langle x|x\rangle + \psi_y\langle x|y\rangle = \psi_x. \tag{1-27}$$

Thus we can write (1-25) as

$$|\Psi\rangle = |x\rangle\langle x|\Psi\rangle + |y\rangle\langle y|\Psi\rangle, \tag{1-28}$$

and (1-26) as

$$|\Psi\rangle = |R\rangle\langle R|\Psi\rangle + |L\rangle\langle L|\Psi\rangle. \tag{1-29}$$

These equations are examples of the *superposition principle;* we can regard any arbitrary polarization as a coherent superposition of, e.g., x and y polarization states, or equivalently as a coherent superposition of right and left circularly polarized states.

Now let us return to the problem of passing a beam through an x polaroid. The classical rules tell us to regard the beam as a superposition of an x polarized beam and a y polarized beam, and that the effect of the polaroid is to throw away the y polarized component and pass only the x polarized component. The absolute value squared of the amplitude of the beam gives us its energy before it passes through, and the absolute square of its x component gives us its energy after it passes through. The fraction of the beam that passes through is given by (1-13). As we have seen, quantum mechanically this fraction gives us the probability of one photon with the initial polarization passing through the polaroid. Written in terms of $|\Psi\rangle$, (1-13) is:

$$\text{Probability} = \frac{|\psi_x|^2}{|\psi_x|^2 + |\psi_y|^2} = |\psi_x|^2 = |\langle x|\Psi\rangle|^2. \tag{1-30}$$

Thus $\langle x|\Psi\rangle$ is the amplitude of the x polarized component of $|\Psi\rangle$, and its absolute value squared is the probability that the photon in the state $|\Psi\rangle$ passes through the x polaroid. We call $\langle x|\Psi\rangle$ the *probability amplitude* for the photon to pass through the x polarizer.

Next consider passing light through a prism that passes right circular but rejects left circular. As we have said, to calculate the probability we write the beam as a coherent sum of right circularly polarized and left circularly polarized light, as in Eq. (1-14) or in terms of the state vectors, as in Eq. (1-29). Then $\langle R|\Psi\rangle$ is the amplitude of the component passed by the prism and, in accord with Eq. (1-15), $|\langle R|\Psi\rangle|^2$ is the probability that a photon in the state $|\Psi\rangle$ will pass through the prism.

The general rule is that if we have a prism that passes only light in the state $|\Phi\rangle$, rejecting light in states orthogonal to $|\Phi\rangle$, then the probability amplitude that a photon in the state $|\Psi\rangle$ will pass through the prism is

$$\langle\Phi|\Psi\rangle, \tag{1-31}$$

and the probability that the photon passes through is

$$|\langle\Phi|\Psi\rangle|^2. \tag{1-32}$$

Notice that this probability is independent of the phase of $|\Psi\rangle$ or $|\Phi\rangle$, though the probability amplitude depends on this phase.

TRANSFORMATION OF BASES

So far, in writing out $|\Psi\rangle$ vectors, as in Eq. (1-17), we have been using the x, y basis. We can equally well use any other basis in writing out the vectors. For example, if we use a basis x', y' related by an angle θ with respect to the old basis, as in Fig. 1-2, then we would write

$$|\Psi\rangle = \begin{pmatrix} \psi_{x'} \\ \psi_{y'} \end{pmatrix} = \begin{pmatrix} \langle x'|\Psi\rangle \\ \langle y'|\Psi\rangle \end{pmatrix}. \tag{1-33}$$

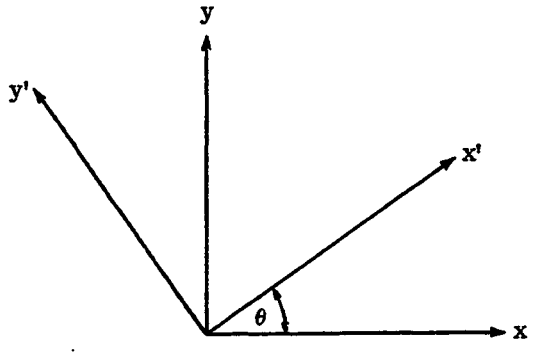

Fig. 1-2
Rotation of plane polarized basis by angle θ.

The problem we want to solve now is how are the components of the vector $|\Psi\rangle$ in the x', y' basis related to the components in the x, y basis. From Eq. (1-25) we can write

$$\langle x'|\Psi\rangle = \langle x'|x\rangle\langle x|\Psi\rangle + \langle x'|y\rangle\langle y|\Psi\rangle \tag{1-34}$$

and

$$\langle y'|\Psi\rangle = \langle y'|x\rangle\langle x|\Psi\rangle + \langle y'|y\rangle\langle y|\Psi\rangle . \tag{1-35}$$

Written in matrix language:

$$\begin{pmatrix} \langle x'|\Psi\rangle \\ \langle y'|\Psi\rangle \end{pmatrix} = \begin{pmatrix} \langle x'|x\rangle & \langle x'|y\rangle \\ \langle y'|x\rangle & \langle y'|y\rangle \end{pmatrix} \begin{pmatrix} \langle x|\Psi\rangle \\ \langle y|\Psi\rangle \end{pmatrix} . \tag{1-36}$$

Thus we know how to transform from the x, y basis to the x', y' basis as soon as we know the matrix in Eq. (1-36); we call this matrix the *transformation matrix* from the x, y basis to the x', y' basis. Notice that Eq. (1-36) is valid for any two bases, not just for the two linear polarized bases we have been considering.

For the particular example that x', y' is the plane polarized basis rotated through an angle θ from the x, y basis, we can write

$$|x\rangle = \cos\theta|x'\rangle - \sin\theta|y'\rangle$$

$$|y\rangle = \sin\theta|x'\rangle + \cos\theta|y'\rangle \tag{1-37}$$

so that

$$\langle x'|x\rangle = \cos\theta, \quad \langle x'|y\rangle = \sin\theta,$$

$$\langle y'|x\rangle = -\sin\theta, \quad \langle y'|y\rangle = \cos\theta, \tag{1-38}$$

and the transformation matrix, which we shall call $\mathcal{R}(\theta)$, is given by

$$\mathcal{R}(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \tag{1-39}$$

Thus

$$\begin{pmatrix} \langle x'|\Psi\rangle \\ \langle y'|\Psi\rangle \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \langle x|\Psi\rangle \\ \langle y|\Psi\rangle \end{pmatrix}. \tag{1-40}$$

We can interpret this equation in either of two ways. First, it tells us the components of the $|\Psi\rangle$ vector in the rotated basis. On the other hand, it is completely equivalent to keep the vector fixed and rotate the basis, or to keep the basis fixed and rotate the vector in the opposite direction. Thus we can regard the vector on the left as a new vector $|\Psi'\rangle$ whose components in the x, y basis are the same as the components of $|\Psi\rangle$ in the x', y' basis, i.e.,

$$\langle x'|\Psi\rangle = \langle x|\Psi'\rangle,$$

$$\langle y'|\Psi\rangle = \langle y|\Psi'\rangle \tag{1-41}$$

When ψ_x and ψ_y are real, the vector $|\Psi'\rangle$ is the vector $|\Psi\rangle$ rotated clockwise by θ, as in Fig. 1-3. Written in terms of $\mathcal{R}(\theta)$,

$$|\Psi'\rangle = \mathcal{R}(\theta)|\Psi\rangle. \tag{1-42}$$

In general, if we transform a vector $|\Psi\rangle$ by $\mathcal{R}(\theta)$, the new components of $|\Psi\rangle$ will bear little resemblance to the old components. Let us ask if there are any state vectors that when operated upon by $\mathcal{R}(\theta)$ are at most multiplied by a constant

$$\mathcal{R}(\theta)|\Psi\rangle = c|\Psi\rangle. \tag{1-43}$$

If a vector satisfies a relation like this it is called an *eigenvector* of $\mathcal{R}(\theta)$, and the number c is called the *eigenvalue* of $\mathcal{R}(\theta)$ belonging to

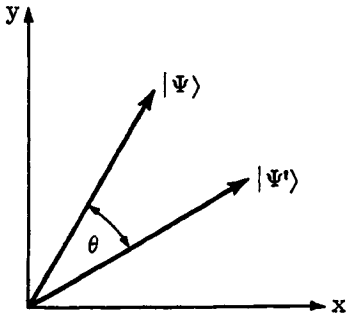

Fig. 1-3
Rotating the basis by angle θ is equivalent to rotating the vectors by angle θ in the opposite direction.

the eigenvector $|\Psi\rangle$. We shall see that the concept of eigenvectors and eigenvalues plays a very important role in quantum mechanics.

To find the eigenvectors and eigenvalues of $\mathcal{R}(\theta)$, let us write it as

$$\mathcal{R}(\theta) = \cos\theta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + i \sin\theta \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{1-44}$$

Let

$$S = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tag{1-45}$$

in the x, y basis. Then we can write the rotation matrix $\mathcal{R}(\theta)$ as

$$\mathcal{R}(\theta) = \cos\theta + iS \sin\theta, \tag{1-46}$$

where the unit matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is understood to multiply the $\cos\theta$.

We first find the eigenvectors and eigenvalues of S and then use Eq. (1-46).

Now if

$$S|\Psi\rangle = \lambda|\Psi\rangle, \tag{1-47}$$

it is easy to see that $\lambda^2 = 1$. This follows directly from the fact that

$$S^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1, \tag{1-48}$$

by the following argument: If $|\Psi\rangle$ satisfies (1-47), then

$$|\Psi\rangle = S^2|\Psi\rangle = S(S|\Psi\rangle) = S(\lambda|\Psi\rangle) = \lambda S|\Psi\rangle = \lambda^2|\Psi\rangle, \tag{1-49}$$

and therefore $\lambda^2 = 1$ and

$$\lambda = \pm 1. \tag{1-50}$$

These are the eigenvalues of S. Now let us find the eigenvectors. The eigenvector corresponding to the eigenvalue +1 satisfies $S|\Psi\rangle = |\Psi\rangle$. We can find the eigenvector by inspection; it is just the vector for right circularly polarized light

$$|R\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix},$$

as may be verified by a trivial calculation. Thus

$$S|R\rangle = |R\rangle \tag{1-51}$$

and similarly we find

$$S|L\rangle = -|L\rangle, \tag{1-52}$$

so that the eigenvector corresponding to the eigenvalue -1 is $|L\rangle$, the vector for left circularly polarized light.

It follows at once that $|R\rangle$ and $|L\rangle$ are also the eigenvectors of $\mathcal{R}(\theta)$ since

$$\mathcal{R}(\theta)|R\rangle = (\cos\theta + i\sin\theta\, S)|R\rangle = (\cos\theta + i\sin\theta)|R\rangle = e^{i\theta}|R\rangle \tag{1-53}$$

and

$$\mathcal{R}(\theta)|L\rangle = (\cos\theta + i\sin\theta S)|L\rangle = (\cos\theta - i\sin\theta)|L\rangle = e^{-i\theta}|L\rangle. \tag{1-54}$$

Thus the components of the vectors for right and left circularly polarized light are changed only by a phase factor under a rotation of the basis. Are $|R\rangle$ and $|L\rangle$ the only eigenvectors of $\mathcal{R}(\theta)$? The answer is yes, as we can see from the fact that we can expand any arbitrary $|\Psi\rangle$ in terms of $|R\rangle$ and $|L\rangle$

$$|\Psi\rangle = |R\rangle\langle R|\Psi\rangle + |L\rangle\langle L|\Psi\rangle.$$

Under a rotation then

$$|\Psi\rangle \rightarrow \mathcal{R}(\theta)|\Psi\rangle = \mathcal{R}(\theta)|R\rangle\langle R|\Psi\rangle + \mathcal{R}(\theta)|L\rangle\langle L|\Psi\rangle = e^{i\theta}|R\rangle\langle R|\Psi\rangle + e^{-i\theta}|L\rangle\langle L|\Psi\rangle. \tag{1-55}$$

Thus we can describe the effect on any $|\Psi\rangle$ of a rotation through angle θ, by saying that the right circularly polarized component is multiplied by a phase factor $e^{i\theta}$ and the left circularly polarized component is multiplied by a phase factor $e^{-i\theta}$. Only if $|\Psi\rangle$ is $|R\rangle$ or $|L\rangle$ is $\mathcal{R}|\Psi\rangle$ a constant times $|\Psi\rangle$.

We shall call S the *spin operator for the photon.* If a photon is right circularly polarized, then its state is an *eigenstate* (which means the same as "eigenvector") of S with eigenvalue +1. We shall say here that such a photon has spin 1. Similarly, left circularly polarized photons have spin −1. Any other state of the photon is a linear superposition of a photon with spin 1 (right circularly polarized) and a photon with spin −1 (left circularly polarized). As seen head on, the electric field vector of right circularly polarized light rotates counterclockwise, while the electric field vector of left circularly polarized light rotates clockwise.

ANGULAR MOMENTUM

The concept of the spin of the photon, which we have introduced in terms of the behavior under rotation has a direct physical meaning; $\hbar$ times the spin of a photon is the component of its *angular momentum* in the z direction. We can see this connection, if we start from

the expression for the angular momentum, $\mathbf{L}$, of the classical electromagnetic field[1]:

$$\mathbf{L} = \frac{1}{4\pi c} \int d^3r\, \mathbf{r} \times [\mathbf{E}(\mathbf{r},t) \times \mathbf{H}(\mathbf{r},t)]. \tag{1-56}$$

Let us consider a plane wave traveling in the z direction. Then $\mathbf{E}$ and $\mathbf{H}$ are in the x, y plane and $\mathbf{L}$ therefore has no component in the z direction. This result is only true for the mythical case of a wave filling all space; for a beam of finite diameter in the x, y plane, $\mathbf{E}\times\mathbf{H}$ must have a component in the x, y plane, near the surface of the beam, in order that the wave satisfy Maxwell's equations.[2] There is actually a z component of the integrand coming from the surface of the beam, and due to the $\mathbf{r}$ in the integrand the total contribution of this term is proportional to the radius of the beam, times the surface area, i.e., proportional to the volume of the beam. We isolate this term simply as follows. When $\nabla\cdot\mathbf{E}(\mathbf{r},t) = 0$, as it is for light, we may introduce the (transverse) vector potential $\mathbf{A}(\mathbf{r},t)$, in terms of which

$$\mathbf{H}(\mathbf{r},t) = \nabla\times\mathbf{A}(\mathbf{r},t), \qquad \mathbf{E}(\mathbf{r},t) = -\frac{1}{c}\frac{\partial\mathbf{A}(\mathbf{r},t)}{\partial t}. \tag{1-57}$$

Then substituting for $\mathbf{H}$ in (1-56), integrating by parts, and assuming the wave to be of finite extent, so that the surface terms at infinity are zero, we find:

$$\mathbf{L} = \frac{1}{4\pi c} \int d^3r\, (\mathbf{E}(\mathbf{r},t)\times\mathbf{A}(\mathbf{r},t)) + \frac{1}{4\pi c}\int d^3r \sum_i E_i(\mathbf{r},t)(\mathbf{r}\times\nabla)A_i(\mathbf{r},t). \tag{1-58}$$

For a plane wave beam traveling in the z direction

$$A_x(\mathbf{r},t) = \frac{c}{\omega} E_x{}^0 \sin(kz-\omega t+\alpha_x),$$

---

[1] See J.D. Jackson, *Classical Electrodynamics* (John Wiley and Sons, New York, 1962).

[2] This is a famous problem; see J.M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons* (Addison-Wesley, Cambridge, 1955), p. 34, for a detailed bibliography on the problem.

$$A_y(\mathbf{r},t) = \frac{c}{\omega} E_y{}^0 \sin(kz - \omega t + \alpha_y) \tag{1-59}$$

except near the edges of the beam. Then the z component of the second term in (1-58) can be neglected. The reason is that $(\mathbf{r}\times\nabla)_z = \partial/\partial\varphi$ [where φ is the polar angle in cylindrical coordinates with axis along $\hat{\mathbf{z}}$] and thus only the surface of the beam contributes; the contribution is proportional to the surface area, not the volume. The z component of the angular momentum of the beam is thus

$$L_z = \frac{1}{4\pi c}\int d^3r\ [E_x(\mathbf{r},t)A_y(\mathbf{r},t) - E_y(\mathbf{r},t)A_x(\mathbf{r},t)],$$

and from (1-59) we find:

$$\begin{aligned} L_z &= \frac{1}{4\pi\omega} E_x{}^0 E_y{}^0 \int d^3r \sin(\alpha_y - \alpha_x) \\ &= \frac{V}{4\pi\omega} E_x{}^0 E_y{}^0 \sin(\alpha_y - \alpha_x). \end{aligned} \tag{1-60}$$

We have been using a real notation in these last few equations. Now let us write (1-60) in terms of the complex $\mathbf{E}$ vector that we have previously been using. Then

$$\begin{aligned} L_z &= \frac{V}{8\pi i\omega} E_x{}^0 E_y{}^0 (e^{i\alpha_y} e^{-i\alpha_x} - e^{-i\alpha_y} e^{i\alpha_x}) \\ &= \frac{V}{8\pi i\omega}(E_x{}^* E_y - E_x E_y{}^*) \\ &= \frac{V}{8\pi\omega}\left(\left|\frac{E_x - iE_y}{\sqrt{2}}\right|^2 - \left|\frac{E_x + iE_y}{\sqrt{2}}\right|^2\right). \end{aligned} \tag{1-61}$$

If we think of the beam as a superposition of a right and a left circularly polarized beam, then

$$\frac{E_x - iE_y}{\sqrt{2}} = E_{RCP}$$

is the right circularly polarized component, and

$$\frac{E_x + iE_y}{\sqrt{2}} = E_{LCP}$$

is the left circularly polarized component. Thus

$$L_z = \frac{V}{8\pi\omega}(|E_{RCP}|^2 - |E_{LCP}|^2). \tag{1-62}$$

Up to here we have been completely classical. Now to make contact with quantum mechanics, we consider a beam consisting of only one photon. Then expressing $\mathbf{E}$ in terms of $|\Psi\rangle$, [Eq. (1-18)], we find

$$L_z = \hbar\left(\left|\frac{\psi_x - i\psi_y}{\sqrt{2}}\right|^2 - \left|\frac{\psi_x + i\psi_y}{\sqrt{2}}\right|^2\right) \tag{1-63}$$

$$= \hbar|\langle R|\Psi\rangle|^2 - \hbar|\langle L|\Psi\rangle|^2.$$

How do we interpret this formula? First of all, it is an experimental fact that if a photon traveling in the z direction is absorbed by matter, then the z component of the angular momentum of the matter either increases by $\hbar$ or decreases by $\hbar$. It never remains the same, nor does it ever change by a value other than $\pm\hbar$. Furthermore, one cannot predict with certainty whether the transfer will be plus or minus $\hbar$. At best we can only predict the probabilities of it being plus or minus $\hbar$. We must interpret Eq. (1-63), therefore, as giving us the *average* value of the angular momentum transfer, averaged over many experiments in which we use a photon in the same polarization state $|\Psi\rangle$ each time. This interpretation is completely analogous to the probabilistic interpretation we gave to the classical formula for the energy that passes through a polaroid filter, for the case that the beam consists of only one photon.

The average value of the angular momentum transfer is $\hbar$ times the probability that the photon transfers angular momentum $\hbar$, plus $-\hbar$ times the probability that the photon transfers angular momentum $-\hbar$. If we then look at Eq. (1-63) for the average value of the angular momentum transfer, we see that $|\langle R|\Psi\rangle|^2$ is the probability that the z component of the angular momentum of the photon is $+\hbar$, and $|\langle L|\Psi\rangle|^2$ is the probability that it is $-\hbar$. In the special case that the photon is right circularly polarized, then $L_z = \hbar$ and we can say that the photon *definitely* has angular momentum $\hbar$. Similarly if the photon is left circularly polarized, then $L_z = -\hbar$, and we can say that the photon *definitely* has angular momentum $-\hbar$. In either case, the z component of the angular momentum is $\hbar$ times the spin value of the photon.

In general, if the photon is neither pure right nor pure left circularly polarized, we can't assign a definite value to its angular momentum; we can only speak of the probability of the angular

momentum of the photon being observed to have value $\hbar$, and the probability of its being observed to have value $-\hbar$. As with the energy, the discreteness of the possible values of the angular momentum leads us into a probabilistic description, if we are to recover the classical description in the limit of a large number of photons.

Let us summarize our description of the z component of the angular momentum of photon. Corresponding to this physical quantity we have an *operator,* i.e., a matrix, $\hbar S$. Photons that are in an *eigenstate,* $|R\rangle$ or $|L\rangle$, of this operator can be assigned a definite value of the z component of angular momentum. This value is just the *eigenvalue* corresponding to the eigenstate that the photon is in. The eigenvalues of $\hbar S$ give the possible values of the z component of angular momentum of the photon. Any other photon state $|\Psi\rangle$ cannot be assigned a definite value of angular momentum, only probabilities for having the possible values. The probability of having a given value is calculated by taking the scalar product of $|\Psi\rangle$ with the eigenstate corresponding to this eigenvalue and then taking the absolute values squared of this product, e.g., $|\langle R|\Psi\rangle|^2$.

This connection between physical quantities and operators, the eigenvalues and the possible values of the physical quantities, and the eigenstates and the probabilities of having one of the possible values, is true in all cases in quantum mechanics, not merely for the particular case we have considered. It is the basis for calculating the expected results of measurements made on quantum mechanical systems.

We notice that the average value of L_z for a given state $|\Psi\rangle$, which is usually called the *expectation value* in that state and is written $\langle L_z\rangle$, is simply given by

$$\langle L_z\rangle = \langle\Psi|\hbar S|\Psi\rangle.$$

To prove this formula we note that from Eq. (1-29)

$$S|\Psi\rangle = S(|R\rangle\langle R|\Psi\rangle + |L\rangle\langle L|\Psi\rangle)$$

$$= |R\rangle\langle R|\Psi\rangle - |L\rangle\langle L|\Psi\rangle,$$

using Eqs. (1-51) and (1-52). Thus

$$\langle\Psi|S|\Psi\rangle = \langle\Psi|R\rangle\langle R|\Psi\rangle - \langle\Psi|L\rangle\langle L|\Psi\rangle, \tag{1-64}$$

so that

$$\hbar\,|\langle R|\Psi\rangle|^2-\hbar\,|\langle L|\Psi\rangle|^2 = \langle\Psi|\hbar S|\Psi\rangle. \tag{1-65}$$

We can make a great advance in notation if we define the *outer product*, $|\Psi\rangle\langle\Phi|$, of two vectors $|\Psi\rangle$ and $|\Phi\rangle$, which is a matrix, as follows

$$|\Psi\rangle\langle\Phi| \equiv \begin{pmatrix}\psi_x\\ \psi_y\end{pmatrix}(\phi_x^* \quad \phi_y^*) = \begin{pmatrix}\psi_x\phi_x^* & \psi_x\phi_y^*\\ \psi_y\phi_x^* & \psi_y\phi_y^*\end{pmatrix}. \tag{1-66}$$

For example,

$$\begin{aligned}
|x\rangle\langle x| &= \begin{pmatrix}1\\0\end{pmatrix}(1 \quad 0) = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\\
|y\rangle\langle y| &= \begin{pmatrix}0\\1\end{pmatrix}(0 \quad 1) = \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\\
|x\rangle\langle y| &= \begin{pmatrix}1\\0\end{pmatrix}(0 \quad 1) = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\\
|y\rangle\langle x| &= \begin{pmatrix}0\\1\end{pmatrix}(1 \quad 0) = \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix},
\end{aligned} \tag{1-67}$$

etc. It is important to notice that if we operate with $|\Psi\rangle\langle\Phi|$ on a vector $|\Theta\rangle$, then the result is the vector $|\Psi\rangle$ times $\langle\Phi|\Theta\rangle$, that is,

$$(|\Psi\rangle\langle\Phi|)|\Theta\rangle = |\Psi\rangle\,(\langle\Phi|\Theta\rangle). \tag{1-68}$$

[One should verify this in detail using Eqs. (1-66) and (1-21).] Thus we needn't write the parentheses in Eq. (1-68), and we can interpret an expression like $|\Psi\rangle\langle\Phi|\Theta\rangle$ in either of the two senses in Eq. (1-68).

Using this outer product notation we can write

$$|x\rangle\langle x| + |y\rangle\langle y| = 1, \tag{1-69}$$

the unit matrix. This equation, which is true for any orthonormal basis, is called the *completeness relation.* Notice that if we operate

with both sides of (1-69) on a vector $|\Psi\rangle$ we find

$$|x\rangle\langle x|\Psi\rangle + |y\rangle\langle y|\Psi\rangle = |\Psi\rangle$$

which we recognize as being simply the expansion, Eq. (1-28), of $|\Psi\rangle$ in terms of the basis vectors $|x\rangle$ and $|y\rangle$.

Now if we multiply the completeness relation in the $|R\rangle, |L\rangle$ basis

$$1 = |R\rangle\langle R| + |L\rangle\langle L| \tag{1-70}$$

on the left by S, we have

$$S = S|R\rangle\langle R| + S|L\rangle\langle L|$$

which, by Eqs. (1-51) and (1-52), becomes

$$\begin{aligned} S &= (+1)|R\rangle\langle R| + (-1)|L\rangle\langle L| \\ &= |R\rangle\langle R| - |L\rangle\langle L|. \end{aligned} \tag{1-71}$$

[This expansion of a matrix in terms of its eigenvalues and eigenvectors will be generally valid for the matrices that represent physical quantities.] Had we known this expression for S before, we could have written down (1-65) immediately.

To end this mathematical interlude, let us note that if we have an arbitrary 2×2 matrix M, which in the x, y basis is explicitly

$$M = \begin{pmatrix} m_{xx} & m_{xy} \\ m_{yx} & m_{yy} \end{pmatrix},$$

then the matrix elements are given by

$$\begin{aligned} m_{xx} &= \langle x|M|x\rangle, \qquad m_{xy} = \langle x|M|y\rangle \\ m_{yx} &= \langle y|M|x\rangle, \qquad m_{yy} = \langle y|M|y\rangle. \end{aligned} \tag{1-72}$$

To see this, we note that, for example,

$$M|x\rangle = \begin{pmatrix} m_{xx} & m_{xy} \\ m_{yx} & m_{yy} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} m_{xx} \\ m_{yx} \end{pmatrix},$$

so that

$$\langle y|M|x\rangle = (0 \quad 1)\begin{pmatrix} m_{xx} \\ m_{yx} \end{pmatrix} = m_{yx},$$

etc. Equation (1-72) can be used for writing down the matrix elements in any basis; for example, in the R, L basis

$$m_{LR} = \langle L|M|R\rangle.$$

AMPLITUDE MECHANICS

In discussing the angular momentum of the photon, we saw that $|\langle R|\Psi\rangle|^2$ was the probability that a photon in state $|\Psi\rangle$ would transfer angular momentum $+\hbar$ in an experiment. Recall that $|\langle R|\Psi\rangle|^2$ also gave us the probability that a photon in state $|\Psi\rangle$ would pass through a prism that passed only right circularly polarized light. In both these cases, $|\langle R|\Psi\rangle|^2$ gives us the probability that the photon in state $|\Psi\rangle$ will behave as if it were in the state $|R\rangle$. The origin of this interpretation lies, of course, in the superposition principle which allowed us to write

$$|\Psi\rangle = |R\rangle\langle R|\Psi\rangle + |L\rangle\langle L|\Psi\rangle.$$

This is one of the strange features of quantum mechanics. Classically a system in a certain state, e.g., a particle at position $\mathbf{x}$, moving with velocity $\mathbf{v}$, never acts as if it were in a different state, say at $\mathbf{x}'$ with velocity $\mathbf{v}'$. But quantum mechanically, because of the superposition principle, a particle in one state, $|\Psi\rangle$, always has a probability of behaving as if it were in state $|\Phi\rangle$. This probability is $|\langle\Phi|\Psi\rangle|^2$, and it vanishes only when $|\Phi\rangle$ and $|\Psi\rangle$ are orthogonal. Only then will a photon in state $|\Psi\rangle$ not have a finite probability of exhibiting all the properties of a photon in state $|\Phi\rangle$.

We shall call $\langle\Phi|\Psi\rangle$ the *probability amplitude* for a photon in the state $|\Psi\rangle$ to be in the state $|\Phi\rangle$. Often one says that $|\langle\Phi|\Psi\rangle|^2$ is the probability for a photon in state $|\Psi\rangle$ *being* in state $|\Phi\rangle$. But, what one really means by this statement is, as we have seen, that $|\langle\Phi|\Psi\rangle|^2$ is the probability of a photon in state $|\Psi\rangle$ behaving in an experiment as if it were in state $|\Phi\rangle$.

We might be tempted to conclude, from measuring the angular momentum transferred to matter, that because the angular momentum transfer is always $+\hbar$ or $-\hbar$, photons were always either in the state $|R\rangle$, with a certain probability, α, or in the state $|L\rangle$ with

probability $1-\alpha$. This is not a correct conclusion, for were it true we could never explain why an x polarized photon *never* passes through a polaroid whose axis lies in the y direction. After all, if

(a) an x polarized photon has probability $|\langle R|x\rangle|^2 = {}^1/_2$ of being right circularly polarized, and a right circularly polarized photon has a probability $|\langle y|R\rangle|^2 = {}^1/_2$ of passing through a y-polaroid, and

(b) an x polarized photon has probability $|\langle L|x\rangle|^2 = {}^1/_2$ of being left circularly polarized, and a left circularly polarized photon has a probability $|\langle y|L\rangle|^2 = {}^1/_2$ of passing through a y-polaroid, then the total probability of an x polarized photon passing through a y-polaroid would be

$$|\langle R|x\rangle|^2|\langle y|R\rangle|^2 + |\langle L|x\rangle|^2|\langle y|L\rangle|^2 = \frac{1}{2}. \tag{1-73}$$

Since x polarized photons *never* pass through y-polaroids, this must be a wrong way of doing the calculation. What we are missing, in thinking of x polarized photons as being either right circularly polarized or left circularly polarized, with equal probability, is the possibility of *interference* between the right and left circularly polarized amplitudes.

The correct calculation of the probability of an x polarized photon passing through a y-polaroid, would take $\langle y|x\rangle$, the probability amplitude for the x polarized photon being in the state $|y\rangle$, and square it to get the probability. Since $\langle y|x\rangle = 0$, this probability is zero. If we regard an x polarized photon as a *superposition* of right and left circularly polarized photons, then the calculation would be the following: The initial state is $|x\rangle$ which can be written as a superposition of $|R\rangle$ and $|L\rangle$:

$$|x\rangle = |R\rangle\langle R|x\rangle + |L\rangle\langle L|x\rangle.$$

Then the probability amplitude for passing through the y-polaroid is

$$\langle y|\Big[|R\rangle\langle R|x\rangle + |L\rangle\langle L|x\rangle\Big] = \langle y|R\rangle\langle R|x\rangle + \langle y|L\rangle\langle L|x\rangle. \tag{1-74}$$

The probability for passing through the polaroid is the absolute square of this amplitude:

$$\begin{aligned}\text{Probability} = |\langle y|R\rangle\langle R|x\rangle + \langle y|L\rangle\langle L|x\rangle|^2 &= |\langle y|R\rangle|^2|\langle R|x\rangle|^2 \\ &+ |\langle y|L\rangle|^2|\langle L|x\rangle|^2 + \langle y|R\rangle\langle R|x\rangle\langle y|L\rangle^*\langle L|x\rangle^* \\ &+ \langle y|R\rangle^*\langle R|x\rangle^*\langle y|L\rangle\langle L|x\rangle.\end{aligned} \tag{1-75}$$

The first two terms on the right give the probability we calculated in (1-73). The true probability is the sum of these two terms plus the last two terms, which represent the *interference* between the right and left circularly polarized photon. It is a trivial calculation to see that the interference terms *exactly* cancel the first two terms, giving a probability zero for an x polarized photon to pass through a y-polaroid; this is, of course, the correct answer.

Let us look more closely at expression (1-74) for the total probability amplitude. $\langle L|x\rangle$ is the probability amplitude for an x polarized photon to be left circularly polarized, and $\langle y|L\rangle$ is the probability amplitude for a left circularly polarized photon to pass through a y-polaroid. Similarly $\langle R|x\rangle$ is the probability amplitude for an x polarized photon to be right circularly polarized, and $\langle y|R\rangle$ is the probability amplitude for a right circularly polarized photon to pass through a y-polaroid. Expression (1-74) then tells us that the total probability amplitude for an x polarized photon to pass through a y-polaroid is the probability *amplitude*, $\langle y|R\rangle\langle R|x\rangle$, for it to pass through as a right circularly polarized photon, plus the probability *amplitude* $\langle y|L\rangle\langle L|x\rangle$ for it to pass through as a left circularly polarized photon. The amplitude for the photon to pass through as a right circularly polarized photon is $\langle L|x\rangle$, the probability amplitude for the x polarized photon to be L polarized, times $\langle y|L\rangle$, the amplitude for an L polarized photon to pass through a y-polaroid. To find the total probability of the photon passing through, we calculate the absolute value squared of the *total* probability amplitude. We do *not* calculate the separate probabilities and then add, as in (1-73).

Notice that the rules for the composition of *probability amplitudes* in quantum mechanics look very much like the classical rules for composition of probabilities:

1. The probability amplitude for two successive possibilities is the product of the amplitudes for the individual possibilities, e.g., the amplitude for the x polarized photon to be right circularly polarized *and* for the right circularly polarized photon to pass through the y-polaroid is $\langle R|x\rangle\langle y|R\rangle$, the product of the individual amplitudes.

2. The amplitude for a process that can take place in one of several *indistinguishable* ways is the sum of the amplitudes for each of the individual ways. For example, the total amplitude for the x polarized photon to pass through the y-polaroid is the sum of the amplitude for it to pass through as a right circularly polarized photon, $\langle y|R\rangle\langle R|x\rangle$, plus the amplitude for it to pass through as a left circularly polarized photon, $\langle y|L\rangle\langle L|x\rangle$, as in Eq. (1-74).

3. The total probability for the process to occur is the absolute value squared of the total amplitude calculated by 1 and 2.

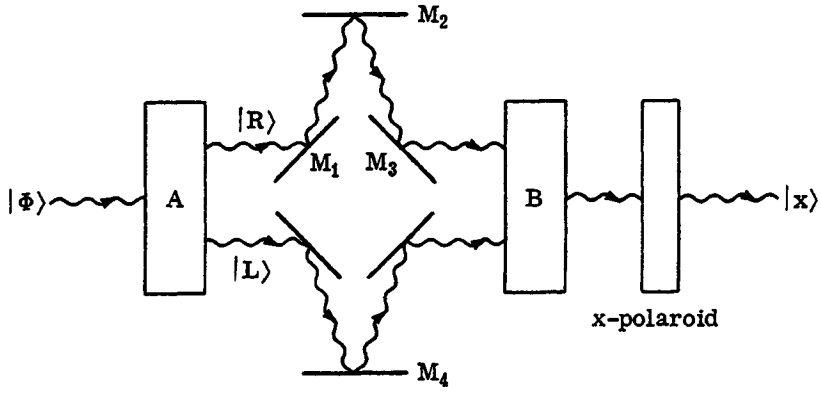

Fig. 1-4

Device for observing whether a photon passes through an x-polaroid as a right or left circularly polarized photon. Prism A splits the beam into right and left circularly polarized components, which are then reflected from mirrors [M_1, M_2, and M_3 for the right circularly polarized component] before being recombined by prism B.

In writing down rule 2 we referred to "indistinguishable ways." This means, for example, that if a photon initially in state $|\Phi\rangle$ passed through a y-polaroid, then just from a knowledge that it passed through, we could not tell if it went through as a right circular or as a left circular photon; these two ways are indistinguishable. If we attempt to distinguish the two possibilities by using some additional measuring device, then the presence of this device so changes the phase relation between the right and left components as to destroy the interference effect. If we had such a device present, and only if it was present, then, in fact, we would observe photons whose initial polarization was in the x direction, passing through the y-polaroid. Such a device is illustrated in Fig. 1-4. A prism A [composed of a birefringent material combined with a quarter wave plate] first splits the light into its right and left circularly polarized components. Then each of these components is passed through a separate sequence of three mirrors, and recombined again by prism B before passing through the polaroid filter. The purpose of the mirrors is to detect the path taken by the photon. A photon of wavelength λ has momentum

$$p = \frac{2\pi\hbar}{\lambda} \tag{1-76}$$

along its direction of motion. Consequently when it is reflected by a mirror, such as M_2, at normal incidence say, it reverses its direction and transfers momentum 2p to the mirror. If a photon passes through the device, and reaches the polaroid, as a right circularly polarized photon, it will cause mirror M_2 to recoil with momentum 2p; if it passes through as a left circularly polarized photon, mirror M_4 recoils. Thus to determine the path taken by the photon we simply observe whether M_2 or M_4 is deflected.

Now to use a mirror, such as M_2 or M_4, as a measuring device we must know its initial momentum p_y in the vertical, or y direction, sufficiently accurately to tell if it receives a momentum transfer 2p from the photon. Thus the initial uncertainty Δp_y in p_y must satisfy

$$\Delta p_y < 2p. \tag{1-77}$$

But from the Heisenberg uncertainty principle,

$$\Delta p_y \Delta y \geq 2\pi\hbar, \tag{1-78}$$

the *position* of the mirror in the y direction must be uncertain by an amount Δy obeying

$$\Delta y \geq \frac{2\pi\hbar}{\Delta p_y} \geq \frac{2\pi\hbar}{2p} = \frac{\lambda}{2}. \tag{1-79}$$

Thus if the y coordinate of the mirror M_2 is uncertain by half a wavelength, the length of the path traveled by the photon in reflecting from M_2 is uncertain by λ [$\lambda/2$ coming and $\lambda/2$ going] — a full wavelength. Consequently the phase of the photon by the time it reaches prism B is *completely* uncertain, and the phase relation between the right and left circularly polarized components is lost.

Mathematically the initial photon state is

$$|\Phi\rangle = |R\rangle\langle R|\Phi\rangle + |L\rangle\langle L|\Phi\rangle; \tag{1-80}$$

let us say that the total phase change of the right circularly polarized component in passing through the device (up to the polaroid) is α_R, and α_L is the total phase change of the left circularly polarized component. Then the photon emerges from prism B in the state

$$|\Phi'\rangle = |R\rangle\langle R|\Phi\rangle\, e^{i\alpha R} + |L\rangle\langle L|\Phi\rangle\, e^{i\alpha L}. \tag{1-81}$$

The amplitude that the photon passes through the x-polaroid is

$$\langle x|R\rangle\langle R|\Phi\rangle\, e^{i\alpha_R} + \langle x|L\rangle\langle L|\Phi\rangle\, e^{i\alpha_L},$$

and the probability is

$$|\langle x|R\rangle|^2|\langle R|\Phi\rangle|^2 + |\langle x|L\rangle|^2|\langle L|\Phi\rangle|^2 + 2\mathrm{Re}[\langle x|R\rangle\langle R|\Phi\rangle\, e^{i\alpha_R}\langle x|L\rangle^*\langle L|\Phi\rangle^* e^{-i\alpha_L}] \tag{1-82}$$

The first two terms are the probabilities that the photon passes through as a right *or* left circularly polarized photon; the last term is the interference. When the mirrors M_2 and M_4 are set up with $\Delta p_y < 2p$ then α_L and α_R are uncertain to within 2π. The observed probability, a result of many measurements, is thus the average of (1-82) over all values of the phases. Since

$$\int_0^{2\pi} \frac{d(\alpha_R - \alpha_L)}{2\pi}\, e^{i(\alpha_R - \alpha_L)} = 0$$

the interference term averages to zero.

Notice that this loss of interference does not depend on our looking at the mirrors; it is due solely to our preparing the mirrors with $\Delta p_y < 2p$ so that they are capable of detecting the photon. If we wished to preserve the interference we must have the positions of all the mirrors fixed to an accuracy $\Delta y < \lambda/2$; this precludes using the mirrors to detect the photons.

UNPOLARIZED LIGHT

All along we have been speaking of photons in definite states of polarization. It is not obvious that we have room in our elaborate machinery for describing *unpolarized* light. The problem is that we have always assumed that we knew the state of polarization of the photon. But consider the following situation. Let us assume that we have a beam of monochromatic light which is composed of photons from two sources, one which always sends out photons in the state $|\Psi_1\rangle$, and one which always sends out photons in the state $|\Psi_2\rangle$. If the sources emit randomly, and are independent, we cannot say whether any particular photon in the beam came from one source or the other, but let us suppose that we know the relative strengths of the two sources. Then we can assign a probability, say p_1, for any photon in the beam to have come from source one, i.e., that it is in state $|\Psi_1\rangle$, and a probability, p_2, for the photon to have come from source two, and thus be in the state $|\Psi_2\rangle$. Of course, $p_1 + p_2 = 1$.

Let us ask how we would write the expectation value, for the photons in the beam, of any physical quantity, say the angular momentum in the direction of propagation. Recall that the average value of L_z was given by $\hbar$ times the probability that a photon transfers angular momentum $+\hbar$, plus $-\hbar$ times the probability that a photon transfers angular momentum $-\hbar$. Now the probability that the photon transfers angular momentum $+\hbar$ is

$$p_1|\langle R|\Psi_1\rangle|^2 + p_2|\langle R|\Psi_2\rangle|^2, \tag{1-83}$$

that is, the probability, p_1, that the photon was in state $|\Psi_1\rangle$, times $|\langle R|\Psi_1\rangle|^2$, the probability that a photon in the state $|\Psi_1\rangle$ will transfer angular momentum $+\hbar$, plus the probability, p_2, that the photon was in state $|\Psi_2\rangle$, times the probability that a photon in state $|\Psi_2\rangle$ will transfer angular momentum $+\hbar$. Similarly, the probability that the photons transfer angular momentum $-\hbar$ is

$$p_1|\langle L|\Psi_1\rangle|^2 + p_2|\langle L|\Psi_2\rangle|^2, \tag{1-84}$$

so that the average value of the angular momentum transfer is

$$\begin{aligned}
&\hbar(p_1|\langle R|\Psi_1\rangle|^2 + p_2|\langle R|\Psi_2\rangle|^2) - \hbar(p_1|\langle L|\Psi_1\rangle|^2 + p_2|\langle L|\Psi_2\rangle|^2) \\
&\quad = p_1(\hbar|\langle R|\Psi_1\rangle|^2 - \hbar|\langle L|\Psi_1\rangle|^2) + p_2(\hbar|\langle R|\Psi_2\rangle|^2 - \hbar|\langle L|\Psi_2\rangle|^2) \qquad (1\text{-}85) \\
&\quad = p_1\langle L_z\rangle_1 + p_2\langle L_z\rangle_2 = p_1\langle\Psi_1|\hbar S|\Psi_1\rangle + p_2\langle\Psi_2|\hbar S|\Psi_2\rangle .
\end{aligned}$$

Thus, the average value of L_z for the beam consisting of two photon states is the average of the expectation value of L_z for state $|\Psi_1\rangle$ and the expectation value of L_z for state $|\Psi_2\rangle$, weighted with the probabilities for finding these states in the beam.

It is very important to realize that the statement "the photon is either in state $|\Psi_1\rangle$ or $|\Psi_2\rangle$, but we don't know which," is *not* the same statement as "the photon is in a state which is a superposition of $|\Psi_1\rangle$ and $|\Psi_2\rangle$." This is the distinction we were emphasizing on p. 21. The point is that when we say that a photon is in a state $|\Phi\rangle$ that is a linear combination of states $|\Psi_1\rangle$ and $|\Psi_2\rangle$,

$$|\Phi\rangle = \alpha|\Psi_1\rangle + \beta|\Psi_2\rangle, \tag{1-86}$$

then we are implying that the relative phase of the coefficients α and β is certain. For example, if $|\Psi_1\rangle = |x\rangle$ and $|\Psi_2\rangle = |y\rangle$, then if $\alpha = 1/\sqrt{2}$ and $\beta = i/\sqrt{2}$, clearly $|\Phi\rangle = |R\rangle$, but if $\alpha = 1/\sqrt{2}$ and $\beta = -i/\sqrt{2}$, then $|\Phi\rangle = |L\rangle$, which is a state completely different from

$|R\rangle$. The relative phase must be fixed to specify the linear combination uniquely. On the other hand, when we say that a photon is either in state $|\Psi_1\rangle$ or state $|\Psi_2\rangle$, but we don't know which, then we are implying that there is no connection between the phases of these states, and hence no possibility of interference effects, which depend critically on delicate phase relations, taking place.

There are really two levels of probability in quantum mechanics. The first, which is called the *pure case*, is when we have a system that is in a definite state (often called a pure state). Then the behavior of this system in a given experiment is governed by the probability amplitude rules we have been discussing. The second case is when the system can be in any of several states, with certain probabilities. This case is called the *mixed case*, and we say that the photon is in a *mixed state*. Then one must calculate the results one expects in a given experiment for each of the separate states that can be present, and take the weighted average of the result over these states, as we have done in Eq. (1-85). In averaging over the various states that can be present, one uses the ordinary classical probability rules; there is no possibility of interference occurring between these states.

If there is complete uncertainty between the phases of α and β in (1-86), then we get the same results as if we had a mixed case in which the photon could be in state $|\Psi_1\rangle$ with probability $p_1 = |\alpha|^2$, and state $|\Psi_2\rangle$ with probability $p_2 = |\beta|^2$. Consider for example, the expectation value of L_Z. For the state $|\Phi\rangle$, (1-86),

$$\begin{aligned}\langle L_Z\rangle = \langle\Phi|\hbar S|\Phi\rangle &= (\alpha^*\langle\Psi_1| + \beta^*\langle\Psi_2|)\hbar S(\alpha|\Psi_1\rangle + \beta|\Psi_2\rangle)\\ &= |\alpha|^2\langle\Psi_1|\hbar S|\Psi_1\rangle + |\beta|^2\langle\Psi_2|\hbar S|\Psi_2\rangle \qquad (1\text{-}87)\\ &\quad + \alpha^*\beta\langle\Psi_1|\hbar S|\Psi_2\rangle + \beta^*\alpha\langle\Psi_2|\hbar S|\Psi_1\rangle .\end{aligned}$$

The first two terms on the right are the classical average of the result for state $|\Psi_1\rangle$ and the result for state $|\Psi_2\rangle$, as in Eq. (1-85). The last two terms represent quantum mechanical interference effects, and depend critically on the relative phase of α and β. If this phase is completely uncertain, and can take on random values for the different photons in the beam, then to get the average value of L_Z, we must average over all possible values of this relative phase. To calculate this average, let us write $\alpha = |\alpha|e^{i\varphi_1}$, $\beta = |\beta|e^{i\varphi_2}$. Clearly neither $|\alpha|^2$ nor $|\beta|^2$ in (1-87) depend on the phases φ_1, φ_2. But

$$\alpha^*\beta = |\alpha\beta|e^{-i(\varphi_1-\varphi_2)},$$

and when we average over the relative phase, $\varphi_1 - \varphi_2$, then $\alpha^*\beta$ and $\beta^*\alpha$ will average to zero, the last two terms in (1-87) vanish, and we are left with the mixed case result (1-85) for the expectation value of L_Z.

We are now ready to answer the question, "what is unpolarized light?" It is light that has equal probability of being in *any* polarization state; the state of unpolarized light is a mixed state.

It turns out that this description of unpolarized light is equivalent to a much simpler one: If $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are *any* two states that form an orthonormal basis, then light that is in a mixed state with equal probability of being in state $|\Psi_1\rangle$ or $|\Psi_2\rangle$ has all the properties of unpolarized light. That is, we could not distinguish experimentally whether unpolarized light was in any state with equal probability, or just in states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ with equal probability.

We can see this as follows. Our first description of unpolarized light calls for us, in calculating any physical quantity, such as $\langle L_Z\rangle$, to average over all polarization states. Now because $|\Psi_1\rangle$ and $|\Psi_2\rangle$ form a basis, we can write any polarization state as a linear combination of them, as in Eq. (1-86). Averaging over all polarization states is equivalent to averaging over all values of α and β, subject only to $|\alpha|^2 + |\beta|^2 = 1$, since $\langle\Phi|\Phi\rangle = 1$.

Let us again consider, as an example, calculating $\langle L_Z\rangle$; but I must emphasize that what we are doing is valid for *any* physical quantity and its corresponding operator. The expectation value of L_Z is given in terms of α, β, $|\Psi_1\rangle$, and $|\Psi_2\rangle$ by Eq. (1-87). In averaging over all possible values of α and β, let us first average over their phases. Then, as we have seen, the last two terms in (1-87) average to zero. Next, we must average over the magnitudes of α and β. Clearly the average of $|\alpha|^2$ will equal the average of $|\beta|^2$ since α and β are completely equivalent. Also, since $|\alpha|^2 + |\beta|^2 = 1$, we see that the average of $|\alpha|^2$ equals the average of $|\beta|^2$ equals $^1/_2$. Thus in carrying out the average for an unpolarized beam we find

$$\langle L_Z\rangle = \frac{1}{2}\langle\Psi_1|\hbar S|\Psi_1\rangle + \frac{1}{2}\langle\Psi_2|\hbar S|\Psi_2\rangle. \tag{1-88}$$

This is just the average we would have found had we assumed the photons to be in either state $|\Psi_1\rangle$ or $|\Psi_2\rangle$ with equal weight.

The average (1-88) for unpolarized light is, of course, independent of the basis $|\Psi_1\rangle$, $|\Psi_2\rangle$, since this was an arbitrary basis. We can see this in another way, if we make use of the completeness relation, e.g. (1-69), for a basis. Suppose $\{|\Phi_1\rangle, |\Phi_2\rangle\}$ is another orthonormal basis. Then we have

$$\langle L_z \rangle = \frac{1}{2}\langle \Psi_1 | \hbar S | (|\Phi_1\rangle\langle\Phi_1| + |\Phi_2\rangle\langle\Phi_2|) | \Psi_1 \rangle$$

$$+ \frac{1}{2}\langle \Psi_2 | \hbar S | (|\Phi_1\rangle\langle\Phi_1| + |\Phi_2\rangle\langle\Phi_2|) | \Psi_2 \rangle$$

$$= \frac{1}{2}(\langle \Phi_1 | \Psi_1 \rangle \langle \Psi_1 | \hbar S | \Phi_1 \rangle + \langle \Phi_1 | \Psi_2 \rangle \langle \Psi_2 | \hbar S | \Phi_1 \rangle)$$

$$+ \frac{1}{2}(\langle \Phi_2 | \Psi_1 \rangle \langle \Psi_1 | \hbar S | \Phi_2 \rangle + \langle \Phi_2 | \Psi_2 \rangle \langle \Psi_2 | \hbar S | \Phi_2 \rangle)$$

$$= \frac{1}{2}\langle \Phi_1 | \hbar S | \Phi_1 \rangle + \frac{1}{2}\langle \Phi_2 | \hbar S | \Phi_2 \rangle ,$$

using the completeness relation for the Ψ basis. Thus unpolarized light is equally well a mixture of equal parts of photons in state $|\Phi_1\rangle$ and photons in state $|\Phi_2\rangle$. For example, the beam that emerges from prism B in Fig. 1-4 is essentially unpolarized when $\Delta y > \lambda/2$ for the mirrors M_4 or M_2, and we do not observe which path the photons took.

BEHAVIOR OF PHOTON POLARIZATION IN MATTER

Up to now we have been considering the effect on photons of prisms and polaroids that pass light in certain polarization states, but reject light in states orthogonal to these. Generally, if we shoot a photon at such a prism it has a certain probability for not emerging from the other side of the prism. Let us consider the passage of photons through crystals that to a first approximation allow all the photons to pass through, but change the state of polarization of the photons.

As an example, consider a "birefringent" crystal, such as calcite. This type of crystal has an axis, called the *optic axis*, with the property that the crystal has a different index of refraction for light polarized parallel to the optic axis than it has for light polarized perpendicular to the optic axis. Let us take a block of calcite with the optic axis in the x, y plane, and pass through photons in the z direction. The photons polarized perpendicular to the optic axis are called the *ordinary* ray, and the photons polarized parallel to the optic axis are called the *extraordinary* ray.

In calcite, for red light, the index of refraction, n_o, for the ordinary ray is about 10% greater than the index, n_e, for the extraordinary

ray. This means that the extraordinary ray travels faster through the crystal than the ordinary ray. Thus for an arbitrarily polarized photon traveling in the z direction, the relative phase of the ordinary and extraordinary components will vary as the photon propagates through the crystal. Consequently the polarization state of the photon when it emerges from the calcite will, in general, be different than its state of polarization on entering the calcite. Let us try to describe this process mathematically.

Let $|o\rangle$ denote the state of polarization of the ordinary ray and let $|e\rangle$ denote the state of polarization of the extraordinary ray; $|e\rangle$ and $|o\rangle$ form an orthonormal basis. Recall that the electric field vector changes by a phase $e^{ikz-i\omega t}$ from point to point in space. Now for a given frequency, ω, the wavenumber k will be different for e and o polarizations:

$$k_e = \frac{\omega n_e}{c}$$

$$k_o = \frac{\omega n_o}{c}. \tag{1-89}$$

If the calcite is of length l in the z direction, then the difference in phase of an e wave between its point of entry in the calcite and its point of departure will be $k_e l$, and similarly an o wave changes phase by $k_o l$. [The time factor $e^{-i\omega t}$ is the same for both polarizations and doesn't effect their relative phase.] The effect of the calcite is to change an incident state $|e\rangle$ into $e^{ike l}|e\rangle$, and an incident $|o\rangle$ state into $e^{iko l}|o\rangle$ as the light passes through the crystal.

Thus, we can describe the effect of the calcite on the polarization state of a photon as follows. Let $|\Psi_{in}\rangle$ denote the polarization state of the photon as it enters the crystal. We can write $|\Psi_{in}\rangle$ as a linear superposition of $|e\rangle$ and $|o\rangle$:

$$|\Psi_{in}\rangle = |e\rangle\langle e|\Psi_{in}\rangle + |o\rangle\langle o|\Psi_{in}\rangle. \tag{1-90}$$

Then the effect of the calcite can be described by multiplying the $|e\rangle$ component by $e^{ike l}$ and the $|o\rangle$ component by $e^{iko l}$; the polarization state $|\Psi_{out}\rangle$ of the photon as it leaves the calcite is therefore

$$|\Psi_{out}\rangle = e^{ike l}|e\rangle\langle e|\Psi_{in}\rangle + e^{iko l}|o\rangle\langle o|\Psi_{in}\rangle = U_l|\Psi_{in}\rangle, \tag{1-91}$$

where

$$U_z = e^{ike z}|e\rangle\langle e| + e^{iko z}|o\rangle\langle o|. \tag{1-92}$$

The probability amplitude that if the photon enters in state $|\Psi\rangle$ it emerges in state $|\Phi\rangle$ (i.e., behaves as if in the state $|\Phi\rangle$) is called the *transition amplitude* from $|\Psi\rangle$ to $|\Phi\rangle$. To calculate this transition amplitude we use the fact that if the photon enters in state $|\Psi\rangle$ it emerges in state $U_l|\Psi\rangle$, and thus the probability amplitude for the emergent photon being in state $|\Phi\rangle$ is $\langle\Phi|U_l|\Psi\rangle$. The square of this amplitude, $|\langle\Phi|U_l|\Psi\rangle|^2$, is usually called the *transition probability* from $|\Psi\rangle$ to $|\Phi\rangle$. It is the probability that if a photon enters the crystal in state $|\Psi\rangle$ it will behave as if it were in the state $|\Phi\rangle$ when it emerges from the crystal. We shall meet the concept of transition probability again and again in quantum mechanics.

Clearly, the state of polarization, $|\Psi_z\rangle$, of the photon after traveling through a length z of the calcite is

$$|\Psi_z\rangle = U_z|\Psi_{in}\rangle, \tag{1-93}$$

It is instructive to consider how $|\Psi_z\rangle$ changes as we move in the z direction. To do this, let us notice that U_z obeys the simple property

$$U_{z+a} = U_a U_z. \tag{1-94}$$

Showing this is left as an exercise. Thus we can determine the polarization vector at the point $z+a$ in terms of the vector at z, by writing

$$|\Psi_{z+a}\rangle = U_{z+a}|\Psi_{in}\rangle = U_a U_z|\Psi_{in}\rangle = U_a|\Psi_z\rangle. \tag{1-95}$$

Now let us suppose that the point $z+a$ is infinitesimally close to z, that is, $k_o a \ll 1$ and $k_e a \ll 1$. Then we can expand the exponentials in the maxtrix U_a, (1-92), and write

$$U_a = (1+ik_o a)|o\rangle\langle o| + (1+ik_e a)|e\rangle\langle e|. \tag{1-96}$$

Note that because $|o\rangle$ and $|e\rangle$ form an orthonormal basis

$$|o\rangle\langle o| + |e\rangle\langle e| = 1.$$

Let us define the wavenumber matrix, K, by

$$K = k_e|e\rangle\langle e| + k_o|o\rangle\langle o|. \tag{1-97}$$

The eigenvalues of K are the possible wavenumbers of photons with frequency ω in the calcite. The eigenvectors are the polarizations that have these wavenumbers. A photon in any other polarization

state can't be assigned a definite wavenumber, but must be regarded as a superposition of the two states $|e\rangle$ and $|o\rangle$ which have definite wavenumbers.

In terms of K we can write U_a, for infinitesimal a, as

$$U_a = 1 + iaK. \tag{1-98}$$

Hence $|\Psi_{z+a}\rangle = (1 + iaK)|\Psi_z\rangle$, or

$$|\Psi_{z+a}\rangle - |\Psi_z\rangle = iaK|\Psi_z\rangle. \tag{1-99}$$

If we divide both sides by a, and let a tend to zero, then the left side becomes the derivative of the vector $|\Psi_z\rangle$ with respect to z, and we have

$$\frac{d}{dz}|\Psi_z\rangle = iK|\Psi_z\rangle. \tag{1-100}$$

This equation, or equivalently Eq. (1-99), tells us how the polarization state changes as we move an infinitesimal distance along the wave. Let us write out Eq. (1-99) in components in the x, y basis:

$$\langle x|\Psi_{z+a}\rangle - \langle x|\Psi_z\rangle = ia\langle x|K|x\rangle\langle x|\Psi_z\rangle + ia\langle x|K|y\rangle\langle y|\Psi_z\rangle. \tag{1-101}$$

This equation says that the *change* in the x component of $|\Psi_z\rangle$ as we move down the beam an infinitesimal amount a, is made up of two pieces; the first proportional to the x component of $|\Psi_z\rangle$, with a constant of proportionality $ia\langle x|K|x\rangle$, and the second proportional to the y component of $|\Psi_z\rangle$ with a constant of proportionality $ia\langle x|K|y\rangle$. Similarly

$$\langle y|\Psi_{z+a}\rangle - \langle y|\Psi_z\rangle = ia\langle y|K|x\rangle\langle x|\Psi_z\rangle + ia\langle y|K|y\rangle\langle y|\Psi_z\rangle. \tag{1-102}$$

The vector $|\Psi_{z+a}\rangle$ is of unit length when $|\Psi_z\rangle$ is; this means that no photons are absorbed between z and z+a. This places a restriction on the matrix K, which we can see by direct calculation. In components

$$\langle\Psi_{z+a}|\Psi_{z+a}\rangle = |\langle x|\Psi_{z+a}\rangle|^2 + |\langle y|\Psi_{z+a}\rangle|^2.$$

Using Eqs. (1-101) and (1-102) and neglecting terms of order a^2, we find

$$\langle \Psi_{z+a}|\Psi_{z+a}\rangle = 1 = |\langle x|\Psi_z\rangle|^2 + |\langle y|\Psi_z\rangle|^2$$
$$+ia(\langle x|K|x\rangle - \langle x|K|x\rangle^*)|\langle x|\Psi_z\rangle|^2$$
$$+ia(\langle y|K|y\rangle - \langle y|K|y\rangle^*)|\langle y|\Psi_z\rangle|^2$$
$$+ia(\langle x|K|y\rangle - \langle y|K|x\rangle^*)\langle y|\Psi_z\rangle\langle x|\Psi_z\rangle^*$$
$$+ia(\langle y|K|x\rangle - \langle x|K|y\rangle^*)\langle x|\Psi_z\rangle\langle y|\Psi_z\rangle^*.$$

Since the first two terms on the right are $\langle \Psi_z|\Psi_z\rangle = 1$, the last four terms must vanish. Since $|\Psi_z\rangle$ is an arbitrary vector, these four terms will vanish only if the coefficients of the components of $|\Psi_z\rangle$ vanish. Thus

$$\langle x|K|x\rangle = \langle x|K|x\rangle^*, \quad \langle y|K|y\rangle = \langle y|K|y\rangle^*$$

$$\langle x|K|y\rangle = \langle y|K|x\rangle^*, \quad \langle y|K|x\rangle = \langle x|K|y\rangle^*. \tag{1-103}$$

That K satisfies these three conditions in fact, can be verified directly from its definition, Eq. (1-97).

If we define the Hermitian adjoint, $K^\dagger$, of K to be the matrix

$$K^\dagger = \begin{pmatrix} \langle x|K|x\rangle^* & \langle y|K|x\rangle^* \\ \langle x|K|y\rangle^* & \langle y|K|y\rangle^* \end{pmatrix}, \tag{1-104}$$

Then (1-103) says that

$$K^\dagger = K; \tag{1-105}$$

K is then said to be Hermitian.

Thus the condition that no photons be absorbed as the beam travels through the crystal implies that the matrix, K, which tells us the infinitesimal change in $|\Psi_z\rangle$ as we move an infinitesimal distance along the beam, must be Hermitian. We shall meet a similar argument again when we come to discuss the Schrödinger equation, which says how state vectors change in time. It is also easy to verify that

$$U_z^\dagger U_z = 1; \tag{1-106}$$

a matrix satisfying this relation is called *unitary*. The fact that U_z is unitary is equivalent to K being Hermitian.

PROBLEMS

1. A man sends a beam of red light along the z axis through a polaroid filter that passes only x polarized light. The beam is initially

polarized at 30° to the x axis, and the total energy content of the beam is quite accurately 10 joules. Estimate the fluctuations in the energy of the beam, i.e., the range of likely energy values, after it passes through the polaroid. How do the fluctuations depend on $\hbar$?

2. (a) Write down a basis corresponding to 45°, 135° polarizations.
(b) Write down a basis that is neither plane nor circularly polarized.

3. (a) Calculate the transformation matrix from the x, y basis to the R, L basis.
(b) Calculate the transformation matrix from the R, L basis to the basis devised in Problem 2(b).
(c) Calculate the transformation matrix from the x, y basis to the basis in Problem 2(b), and show that it is the product of the matrix calculated in Problem 3(b) and the matrix calculated in Problem 3(a). In which order must you multiply these matrices?
(d) Show in general that the product of
   (i) the transformation matrix from a basis, 1, to a basis, 2, with
   (ii) the transformation matrix from basis 2 to a third basis, 3,

   is the transformation matrix from basis 1 to basis 3. Hint: make extensive use of the completeness relation, e.g., (1-69), (1-70), for a basis.

4. (a) Show that $\langle\Phi|M|\Psi\rangle^* = \langle\Psi|M^\dagger|\Phi\rangle$, for any $|\Phi\rangle$ and $|\Psi\rangle$.
(b) Show that if $M|\Psi\rangle = \lambda|\Psi\rangle$ then $\langle\Psi|M^\dagger = \lambda^*\langle\Psi|$.

5. Show that the transformation matrix from one basis to another is unitary.

6. (a) Show that the matrix $|\Phi\rangle\langle\Phi|$ is Hermitian.
(b) Show that the photon spin operator S is Hermitian. Generally all physical quantities are represented by Hermitian matrices.

7. Let $|x'(\theta)\rangle$, $|y'(\theta)\rangle$ denote the basis tilted at an angle θ to the x, y basis. Show that the components of a vector $|\Psi\rangle$ in this basis

$$\langle x'(\theta)|\Psi\rangle, \qquad \langle y'(\theta)|\Psi\rangle$$

obey the differential equations

$$-i\frac{\partial}{\partial\theta}\langle x'(\theta)|\Psi\rangle = \langle x'(\theta)|S|\Psi\rangle$$

$$-i\frac{\partial}{\partial\theta}\langle y'(\theta)|\Psi\rangle = \langle y'(\theta)|S|\Psi\rangle.$$

Solve these equations explicitly for $|\Psi\rangle = |R\rangle$ and $|\Psi\rangle = |L\rangle$.

8. The probability that a photon in state $|\Psi\rangle$ passes through an x-polaroid is the average value of a physical observable which might be called the "x-polarizedness." Write down the operator, P_x, corresponding to this observable. Show that it is Hermitian. What are its eigenvalues and eigenstates? Write down its representation in terms of its eigenvalues and eigenstates [as in (1-71)]. Verify that the probability that a photon in state $|\Psi\rangle$ passes through the x-polaroid is $\langle\Psi|P_x|\Psi\rangle$.
9. Photons polarized at 30° to the x axis are sent through a y-polaroid. An attempt is made to determine how frequently the photons that pass through the polaroid, pass through "as right circularly polarized photons," and how frequently they pass through "as left circularly polarized photons"; this attempt is made as follows:
   First, a prism that passes only right circular polarized light is placed between the source of the 30° polarized photons and the y-polaroid, and it is determined how frequently the 30° photons pass through the y-polaroid. Then this experiment is repeated with a prism that passes only left circular polarized light instead of the one that passes only right. Show by explicit calculation that the sum of the probabilities for passing through the y-polaroid measured in these two experiments is different from the probability that one would measure if there were *no* prism in the path of the photon and only the y-polaroid.
   Relate this experiment to the two-slit diffraction experiment.
10. A beam with a certain number of photons is prepared by reflecting from a mirror for a moment a monochromatic beam from a laser; such a laser beam has, for our purposes, a very well defined phase. The number of photons in the reflected beam is determined by measuring the momentum transferred to the mirror in the reflection. Show that ΔN, the uncertainty in the number of photons in the reflected beam, times $\Delta\varphi$, the uncertainty in the phase of the reflected beam is greater than 2π. Show that this uncertainty relation is independent of the angle of incidence of the laser beam on the mirror. The uncertainty relation $\Delta N \Delta\varphi \geq 2\pi$ between the number of photons in a wave and its phase is quite generally true, and doesn't depend on how the beam was prepared.
11. A photon polarized at an angle θ to the optic axis is sent in the z direction through a slab of calcite 10^{-2} cm thick in the z direction. Assume the optic axis to lie in the x, y plane. Calculate, as a function of θ, the transition probability for the photon to emerge left circularly polarized. Sketch the result. Let the

frequency of the light be given by $c/\omega = 5000$ Å, and let $n_e = 1.50$ and $n_o = 1.65$ for the calcite.

12. Using calcite and polaroid, devise a filter that will pass light of frequency $c/\omega = 5000$ Å only if it is right circularly polarized. Use the same indices as in 11.
13. What is the condition on the length of the calcite that, for frequency ω, $|\Psi_{out}\rangle$ is *always*, to within a phase factor, the same state as $|\Psi_{in}\rangle$.
14. Turpentine is an "optically active" substance. If we shoot plane polarized light into turpentine then it emerges with its plane of polarization rotated. Specifically, turpentine induces a left-hand rotation of about 5° per cm of turpentine that the light traverses. Write down the transition matrix that relates the incident polarization to the emergent polarization. Show that the matrix is unitary. Find its eigenvectors and eigenvalues, as a function of the length of turpentine traversed.
15. Unpolarized light traveling in the z direction is sent through a block of calcite whose optic axis lies in the x, y plane. What is the effect of the calcite on the polarization properties of the beam? What will turpentine do to an unpolarized beam?
16. The *trace* of an $n \times n$ matrix is defined as the sum of its diagonal components, i.e., if

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \\ \vdots & \vdots & \ddots & \\ a_{n1} & & & a_{nn} \end{pmatrix}$$

then

$$\text{tr } A = a_{11} + a_{22} + a_{33} + \ldots + a_{nn}.$$

Show that the trace of a matrix is independent of the basis in which the matrix is written, i.e., that it is invariant under a unitary transformation of the matrix,

$$A \rightarrow U^\dagger A U,$$

where U is unitary. Show that $\text{tr } AB = \text{tr } BA$.

17. (a) Suppose that a photon is in the state $|\Psi\rangle$. Let

$$P_\Psi = |\Psi\rangle\langle\Psi|.$$

Show that the expectation value for the photon of a physical quantity represented by the operator Q is tr $P_\Psi Q$. The matrix P_Ψ is called the *density matrix* for the pure state $|\Psi\rangle$.

(b) Suppose now that the photon is in a state that is a mixture of state $|\Psi_1\rangle$ with probability p_1, $|\Psi_2\rangle$ with probability p_2, etc. where $\sum_i p_i = 1$.

Let

$$\rho = \sum_i p_i |\Psi_i\rangle \langle \Psi_i|.$$

Show that the expectation value for this mixed state of a physical quantity Q is

$$\langle Q \rangle = \mathrm{tr}\, \rho Q.$$

The matrix ρ is called the *density matrix* for the mixed state.

(c) Show tr $\rho = 1$.

18. (a) The expectation value of any observable, Q, can be calculated in terms of the density matrix as tr ρQ. The density matrix thus contains all information available about the state, pure or mixed, of a photon. Suppose that we have a beam of photons whose polarization state is unknown to us; it might be pure or mixed. We would like to perform several measurements that will completely determine the state. What measurements will determine the polarization state, and what is the minimum number necessary?

(b) What is ρ for unpolarized light? What measurements would determine that the light was unpolarized?

(c) Find ρ for a *mixed* state consisting of 50% x polarized light and 50% right circularly polarized light. Find two orthogonal states that give the same density matrix.

Chapter 2
NEUTRAL K MESONS

The formalism we have developed to discuss the polarization of the photon is applicable to many other problems. We shall consider, as an example, neutral K mesons. These mesons are produced in strong interaction processes like

$$\pi^- + p \rightarrow \Lambda^0 + K^0. \tag{2-1}$$

Notice that charge is conserved in this reaction; the left side has total charge zero, and so does the right side. Similarly, there exists another "quantum number," called *strangeness,* that we can assign to the particles that participate in strong interactions, and it is an experimental fact that strangeness is conserved in strong interactions. Protons, neutrons, and π mesons have strangeness zero, whereas Λ^0 particles have strangeness -1, and K^0 mesons have strangeness $+1$. The left side of (2-1) has total strangeness zero, and so does the right side.

Corresponding to the K^0 meson, there is an antiparticle, called the $\overline{K}^0$ meson. It must have the opposite charge and strangeness of the K^0 meson, and thus it is neutral and has strangeness -1. A typical reaction for a $\overline{K}^0$ is the absorption process.

$$\overline{K}^0 + p \rightarrow \Lambda^0 + \pi^+. \tag{2-2}$$

[Verify that strangeness is conserved in this strong interaction.]

One of the interesting facts about neutral K mesons is that they can be in a linear superposition of K^0 and $\overline{K}^0$ states, exactly as photons can be in a linear superposition of right and left circularly polarized states. Let us represent the states of neutral K mesons by two-dimensional complex vectors, the same type of vectors we used to represent photon polarization. Thus we let $|K^0\rangle$ denote the state

in which the meson is a K^0, and we let $|\overline{K}^0\rangle$ denote the state in which the meson is a $\overline{K}^0$. These states are orthogonal, and we choose them to be normalized to one. Thus $|K^0\rangle$ and $|\overline{K}^0\rangle$ form an orthonormal basis. The general state of a neutral K meson is a linear combination of $|K^0\rangle$ and $|\overline{K}^0\rangle$. These basis states are independent of time; we will put all time dependence in the coefficients in this linear combination. One may think of $|K^0\rangle$ as analogous to the state $|R\rangle$ and $|\overline{K}^0\rangle$ as analogous to $|L\rangle$.

We can define a strangeness operator, S, for K mesons, by writing

$$S|K^0\rangle = |K^0\rangle, \qquad S|\overline{K}^0\rangle = -|\overline{K}^0\rangle. \tag{2-3}$$

Thus in the $|K^0\rangle$, $|\overline{K}^0\rangle$ basis

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2-4}$$

The eigenvectors of S are $|K^0\rangle$ and $|\overline{K}^0\rangle$, the states that have a definite strangeness, and the corresponding eigenvalues are the strangeness values of the K^0 and $\overline{K}^0$ mesons. Another way of writing S that doesn't involve writing out the components explicitly is

$$S = |K^0\rangle\langle K^0| - |\overline{K}^0\rangle\langle\overline{K}^0|, \tag{2-5}$$

which we find from operating with S on both sides of the completeness relation for the K^0, $\overline{K}^0$ basis.

We can also define a *charge conjugation* operation for neutral K mesons that changes particles into antiparticles and vice versa. This operation, denoted by CP, is defined by

$$CP|K^0\rangle = |\overline{K}^0\rangle, \qquad CP|\overline{K}^0\rangle = |K^0\rangle. \tag{2-6}$$

In the K^0, $\overline{K}^0$ basis we have

$$CP = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2-7}$$

The eigenstates of CP are

$$|K_S\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle + |\overline{K}^0\rangle), \qquad |K_L\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle - |\overline{K}^0\rangle), \tag{2-8}$$

as may be verified directly from (2-6):

$$CP|K_S\rangle = |K_S\rangle, \quad CP|K_L\rangle = -|K_L\rangle. \tag{2-9}$$

The eigenvalues of CP are +1 and −1, and the corresponding physical quantity is called *charge conjugation parity*. Thus a neutral K meson in the $|K_S\rangle$ state has positive charge conjugation parity, while a neutral K meson in the $|K_L\rangle$ state has negative charge conjugation parity. We cannot assign a definite value of strangeness to the K_S and K_L states; they are linear combinations, Eq. (2-8), of a state with strangeness +1 and a state with strangeness −1.

It is trivial to verify that $|K_S\rangle$ and $|K_L\rangle$ form an orthonormal basis. We can express $|K^0\rangle$ and $|\overline{K}^0\rangle$ in terms of them as

$$|K^0\rangle = \frac{1}{\sqrt{2}}(|K_S\rangle + |K_L\rangle), \quad |\overline{K}^0\rangle = \frac{1}{\sqrt{2}}(|K_S\rangle - |K_L\rangle). \tag{2-10}$$

We cannot assign a definite value of charge conjugation parity to the $|K^0\rangle$ and $|\overline{K}^0\rangle$ states; they are linear combinations of a state with CP = +1 and a state with CP = −1.

The neutral K mesons decay by weak interactions. As strangeness is conserved in strong interactions, charge conjugation parity is conserved in weak interactions. [Actually, CP is not perfectly conserved in weak interactions, but we shall only examine the consequences of exact conservation; see Problem 2.] One sees two kinds of decays of neutral K mesons. The first type of decay is into a $\pi^+ + \pi^-$ or $2\pi^0$ state that has CP = +1. Because CP is conserved, the state before the decay must have had CP = +1, and thus the state that decays into $\pi^+ + \pi^-$ or $2\pi^0$s must be K_S,

$$K_S \to \pi^+ + \pi^-, \quad K_S \to \pi^0 + \pi^0. \tag{2-11}$$

These processes occur in a time $\tau_S \approx 0.9 \times 10^{-10}$ sec. The other type of decay is into states with CP = −1, such as $\pi\pi\pi$, $\pi\mu\nu$, and $\pi e\nu$. The K state before the decay must have been, by CP conservation, a K_L state:

$$K_L \to \pi\pi\pi, \quad K_L \to \pi e\nu, \quad K_L \to \pi\mu\nu. \tag{2-12}$$

These processes occur in a time $\tau_L \approx 518 \times 10^{-10}$ sec, a very much longer time than that for K_S decay. [The subscripts S and L stand for short- and long-lived.]

Let us consider how the states of neutral K mesons change in time. First recall that when we discussed calcite in Chapter 1, we

found that a photon with frequency ω in the $|o\rangle$ state had a definite wavenumber, $\omega n_o/c$, a photon with frequency ω in the $|e\rangle$ state also had a definite wavenumber, $\omega n_e/c$, but that we couldn't assign a definite wavenumber to a photon that was in a linear combination of $|o\rangle$ and $|e\rangle$ states. A very similar situation occurs with neutral K mesons. If the meson is short lived then its state changes in time with a definite frequency,

$$\omega_S = \frac{E_S}{\hbar} \tag{2-13}$$

where

$$E_S = (p^2c^2 + m_S{}^2c^4)^{1/2} \tag{2-14}$$

is the energy of the K_S meson, m_S is its mass, and p is its momentum. If the meson is long lived then its state changes in time with a frequency

$$\omega_L = \frac{E_L}{\hbar} \tag{2-15}$$

where

$$E_L = (p^2c^2 + m_L{}^2c^4)^{1/2}; \tag{2-16}$$

m_L is the mass of the K_L meson. [This connection between frequency, i.e., the time rate of change of the phase, and the energy, is a fundamental law of quantum mechanics.] If the state of the K is a linear combination of K_S and K_L, such as K^0, then it doesn't have a single frequency; rather it varies as a sum of two frequencies, one for its K_S component, and one for its K_L component.

Suppose that at t = 0, the state $|\Psi(t)\rangle$ of a neutral K meson is pure K_S

$$|\Psi(t=0)\rangle = |K_S\rangle. \tag{2-17}$$

We expect that at a later time, the probability for finding the particle in this state should decrease by a factor e^{-t/τ_S}, because of the exponential decay law for the decay of the K_S into $\pi\pi$. Now this probability is simply $|\langle K_S|\Psi(t)\rangle|^2$. Thus the amplitude $\langle K_S|\Psi(t)\rangle$ must vary in time with a factor $e^{-t/2\tau_S}$, and the state $|\Psi(t)\rangle$ must also have this time dependence, since the basis state $|K_S\rangle$ given by (2-8) is constant in time. Furthermore, as we have discussed, the state

$|\Psi(t)\rangle$ varies in time by a phase factor $e^{-i\omega_S t}$, where $\hbar\omega_S$ is the energy of the K_S meson. Thus putting these two factors together we have

$$|\Psi(t)\rangle = e^{-i\omega_S t - t/2\tau_S}|K_S\rangle \tag{2-18}$$

as the state of the K_S after time t. Similarly, if we start out with a K meson that is in a pure K_L state at t = 0, then at time t its state will be

$$|\Psi(t)\rangle = e^{-i\omega_L t - t/2\tau_L}|K_L\rangle. \tag{2-19}$$

QUANTUM INTERFERENCE EFFECTS

Some rather striking quantum mechanical effects can occur with K mesons. As a first example, we consider how K^0 particles can turn into their antiparticles, $\overline{K}^0$. Suppose that at t = 0 we produce, by process (2-1) for example, a neutral K meson in the $|K^0\rangle$ state. [We know that if strangeness is conserved in strong interactions the neutral K produced must have a definite strangeness value, +1, and hence be a K^0.] Let us consider how the state $|\Psi(t)\rangle$ of this meson changes in time. At t = 0,

$$|\Psi(t=0)\rangle = |K^0\rangle = \frac{1}{\sqrt{2}}(|K_S\rangle + |K_L\rangle). \tag{2-20}$$

The change of this state in time is governed by the change of its K_S and K_L components in time; these are the components that have a well-defined frequency. From (2-18) and (2-19) we find that at time t,

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}}\left[e^{-i\omega_S t - t/2\tau_S}|K_S\rangle + e^{-i\omega_L t - t/2\tau_L}|K_L\rangle\right]. \tag{2-21}$$

The probability amplitude that the meson in this state is a $\overline{K}^0$ at time t is $\langle\overline{K}^0|\Psi(t)\rangle$. From Eq. (2-8), we find $\langle\overline{K}^0|K_S\rangle = -\langle\overline{K}^0|K_L\rangle = 1/\sqrt{2}$. Thus

$$\langle\overline{K}^0|\Psi(t)\rangle = \frac{1}{2}\left(e^{-i\omega_S t - t/2\tau_S} - e^{-i\omega_L t - t/2\tau_L}\right). \tag{2-22}$$

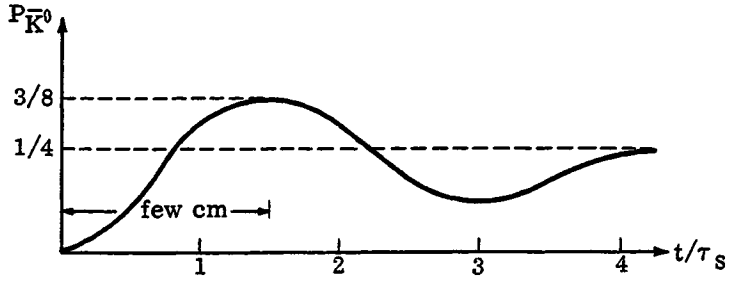

Fig. 2-1

Probability of observing a $\bar{K}^0$ at time t, when the state at time 0 is a K^0. Time is measured in units of τ_S.

The probability, then, that the meson in state $|\Psi(t)\rangle$ will behave as a $\bar{K}^0$ is

$$P_{\bar{K}^0}(t) = |\langle \bar{K}^0 | \Psi(t) \rangle|^2$$

$$= \frac{1}{4}\left[e^{-t/\tau_S} + e^{-t/\tau_L} - 2e^{-t(\tau_S^{-1}+\tau_L^{-1})/2} \cos(\omega_S - \omega_L)t \right]. \tag{2-23}$$

The frequency difference $\omega_S - \omega_L$ is in order of magnitude just $c^2/\hbar$ times the mass difference, $m_S - m_L$, and experimentally $m_L - m_S \approx \hbar/2c^2\tau_S$. Thus the plot of $P_{\bar{K}^0}(t)$ looks something like Fig. 2-1 for $t \ll \tau_L$. It is perhaps more useful to think of the abcissa of Fig. 2-1 as being the distance $l = vt$, that the K meson has traveled from its point of creation; v is its velocity. The picture then shows about the first 10 cm after the creation. Experimentally one observes the creation of the K^0 by looking for the Λ^0 (or rather its decay products) in (2-1). When the K^0 is created, it has no amplitude for being a $\bar{K}^0$, since $\langle \bar{K}^0 | K^0 \rangle = 0$. However, as the particle travels along, its K_S component decays away, leaving just K_L, which contains a large $\bar{K}^0$ component. Of course, after a very large distance, $\sim 10^2$ meters [$= p\tau_L/m_L$ where p is the *momentum* of the meson], the K_L component will also have decayed, but that is too far away to show on Fig. 2-1. The $\bar{K}^0$ component is observed when it undergoes a reaction like (2-2); one looks for the Λ^0 that is created in this reaction. The wiggle in the curve is a result of the $\cos(\omega_S - \omega_L)t$ term in (2-23). This term is due to quantum mechanical interference between the K_S and K_L amplitudes of the particle, and it provides an experimental determination of the mass difference $|m_S - m_L|$.

A good optical analogy to this effect would be a situation in which we had a crystal that absorbed x polarized light strongly, but y polarized light only weakly. Then if we shined right circular light through this crystal, after a small distance there would be a large amplitude for the light to be left circularly polarized.

Another effect of the same nature is the conversion of K_L mesons into K_S mesons. If we produce a K beam, then after a few centimeters the K_S component will decay away, and the beam will be pure K_L. Any decays observed then will be of the type in (2-12). Now suppose that we pass this K_L beam through a thin slab of matter. The K mesons will undergo strong interactions with the protons and neutrons in the matter. The K^0 component of the K_L can be scattered, possibly out of the beam, while the $\overline{K}^0$ component can be both scattered and also absorbed via the reaction (2-2).

Mathematically, the effect of the matter is to multiply the $|K^0\rangle$ component of the beam by a factor $\alpha e^{i\varphi}$, and the $|\overline{K}^0\rangle$ component by a factor $\beta e^{i\varphi'}$. The numbers α and β are positive and generally will be smaller than one, representing loss of particles from the beam due to scattering or absorption. The phase shifts φ and φ' are real. Thus if the state of the beam before the absorber is pure K_L

$$|\Psi_{\text{before}}\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle - |\overline{K}^0\rangle), \tag{2-24}$$

then after the absorber it will be

$$|\Psi_{\text{after}}\rangle = \frac{1}{\sqrt{2}}(\alpha e^{i\varphi}|K^0\rangle - \beta e^{i\varphi'}|\overline{K}^0\rangle). \tag{2-25}$$

The amplitude for the beam to be in the K_S state is therefore

$$\langle K_S|\Psi_{\text{after}}\rangle = \frac{1}{2}(\alpha e^{i\varphi} - \beta e^{i\varphi'}), \tag{2-26}$$

which is nonzero if $\alpha \neq \beta$ or $\varphi \neq \varphi'$.

Thus the piece of matter has the effect of transforming the pure K_L beam into a beam with a K_S component, and it then becomes possible to see K_S decays (2-11) on the far side of the matter. This effect, which is well established experimentally, is called the *regeneration* of K_S mesons; after all the K_S's in the initial beam have decayed away, more can be made by passing the pure K_L beam through matter.

PROBLEMS

1. Suppose that a pure K_L beam is sent through a thin absorber whose only effect is to change the relative phase of the K^0 and $\overline{K}^0$ amplitudes by 10°. Calculate the number of K_S decays, relative to the incident number of particles, that will be observed in the first 5 cm beyond the absorber. Assume for simplicity that the particles have momentum = mc.
2. If CP is not conserved in the decay of neutral K mesons, then the states of definite energy are no longer the K_L, K_S states, but are slightly different states $|K_L'\rangle$ and $|K_S'\rangle$. One can write for example $|K_L'\rangle \sim (1+\varepsilon)|K^0\rangle - (1-\varepsilon)|\overline{K}^0\rangle$ where ε is a very small complex number ($|\varepsilon| \sim 2 \times 10^{-3}$) that is a measure of the lack of CP conservation in the decays. The amplitude for a particle to be in $|K_L'\rangle$ (or $|K_S'\rangle$) varies as $e^{-i\omega_L t - t/2\tau_L}$ (or $e^{-i\omega_S t - t/2\tau_S}$) where $\hbar\omega_L = (p^2c^2 + m_L{}^2c^4)^{1/2}$, etc. As before $\tau_L \gg \tau_S$.
   (a) Write out normalized expressions for the states $|K_S'\rangle$ and $|K_L'\rangle$ in terms of $|K^0\rangle$ and $|\overline{K}^0\rangle$.
   (b) Calculate the ratio of (i) the amplitude for a long-lived K to decay into two pions (a CP = +1 state) to (ii) the amplitude for a short-lived K to decay into two pions. What does a measurement of the ratio of these decay rates tell one about ε?
   (c) Suppose that a beam of purely long-lived K mesons is sent through an absorber whose only effect is to change the relative phase of the K^0 and $\overline{K}^0$ components by δ. Derive an expression for the number of two pion events observed as a function of the time of travel from the absorber. How well would such a measurement (given δ) enable one to determine the phase of ε and the short-long mass difference?

Chapter 3
THE MOTION OF PARTICLES IN QUANTUM MECHANICS

The motion of a particle in quantum mechanics is described by a (complex) wave function, $\psi(\mathbf{r}, t)$, that gives the probability amplitude for finding the particle at point $\mathbf{r}$ at time t. The absolute value squared $|\psi(\mathbf{r}, t)|^2$ of the wave function times a volume element $d^3\mathbf{r}$ is the probability of finding the particle at time t in the volume element $d^3\mathbf{r}$ about t. Because the wave $\psi(\mathbf{r}, t)$ is a probability amplitude, it doesn't tell us how any one particle *will* behave, but rather it tells us the behavior of a large statistical sample of particles subjected to identical conditions. We found the same situation in discussing the polarization state of the photon; we could not say with certainty how any one photon would behave when, for example, passed through a polaroid. We could only give the fraction of a large number of identical photons that passed through the polaroid, and hence only the probability that any one would pass through.

THE SCHRÖDINGER EQUATION[1]

Suppose that we know the wave function for a particle at a certain time t. How will it change over the course of time? Let us begin by answering this question for a particle that moves in one dimension only. It will be most convenient for us to divide the line along which the particle moves into very small intervals, each of length λ (Fig. 3-1). We

[1] This discussion, similar to material in *The Feynman Lectures in Physics*, Vol. III, should not be regarded as a derivation of the Schrödinger equation; rather, it is an attempt to dissect it to see how it works.

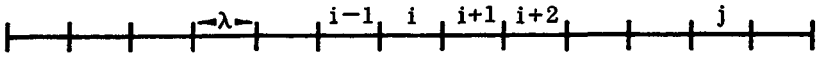

Fig. 3-1
One dimension divided into intervals of length λ.

label the intervals by letters, i, j, ..., and let $\psi_i(t)$ be the probability amplitude for finding the particle in the interval i at time t. Then $|\psi_i(t)|^2$ is the probability of finding the particle in the interval i at time t_i and since the particle must be somewhere along the line, the total probability, summed over all the intervals must be one,

$$\sum_i |\psi_i(t)|^2 = 1. \tag{3-1}$$

The amplitudes $\psi_i(t)$ are like the components of a giant vector

$$|\Psi\rangle = \begin{pmatrix} \psi_1(t) \\ \psi_2(t) \\ \vdots \\ \psi_{i-1}(t) \\ \psi_i(t) \\ \psi_{i+1}(t) \\ \vdots \end{pmatrix} \tag{3-2}$$

very analogous to the state vector that described the polarization state of the photon. The photon vector had only two components, since there were only two independent possible polarizations, e.g., every vector could be written as a superposition of x and y polarization vectors. On the other hand, this vector (3-2), has an *infinite* number of components, since there are an infinite number of different intervals along the line in which the particle can be.

Let us suppose that at time t, there is some amplitude for the particle to be in the interval i. Then because the particle is free to move about, this amplitude will "leak" into the neighboring intervals,

exactly as a drop of oil on a table will slowly spread out over the table. Let us ask how much amplitude flows into the interval i + 1 from i in a time Δt. The motion of the particle from i to i + 1 is described by a probability amplitude, in very much the same way as we described the passage of photon through a polaroid by a probability amplitude [e.g., Eq. (1-74)]. The increase in the probability amplitude in i + 1 is just the probability amplitude for the particle if it is in i to move to i + 1 in time Δt times the probability amplitude, $\psi_i(t)$, that it is in i to begin with. The probability amplitude for the particle to move from i to i + 1 in time Δt will, for sufficiently small Δt, be proportional to Δt. Let us call the constant of proportionality $-iw_{i+1,i}$, where the i before the w stands for $\sqrt{-1}$. Thus the total amount of amplitude that flows from i to i + 1 in time Δt will be

$$-i\Delta t w_{i+1,i}\,\psi_i(t). \tag{3-3}$$

The amplitude in interval i at time t + Δt will be increased by amplitude flowing in from the neighboring intervals, and also decreased by amplitude flowing out. We can write the amplitude $\psi_i(t + \Delta t)$ as

$$\psi_i(t+\Delta t) = \psi_i(t) - i\Delta t w_{i,i-1}\psi_{i-1}(t) - i\Delta t w_{i,i+1}\psi_{i+1}(t) - i\Delta t w_{i\,i}\psi_i(t). \tag{3-4}$$

The first term on the right is the amplitude that was in i at time t. The second term represents the amplitude that has flowed in from i − 1, and the third term represents the amplitude that has flowed in from i + 1. If Δt is sufficiently small, we needn't worry about amplitude flowing into i from intervals further away than the nearest neighbors. The last term in (3-4) includes the decrease in amplitude in i due to its flowing out; this is proportional to Δt and to $\psi_i(t)$. There is another possible change in the amplitude $\psi_i(t)$ that is included in this last term. Even if there were no flow from interval to interval, the amplitude in interval i could change in time Δt by a phase factor of absolute value one, since this wouldn't change the probability of the particle being in the interval i. We shall return to this point shortly.

The w coefficients are not completely arbitrary, but are restricted by the fact that the total probability for finding the particle somewhere must remain one at all times. Thus

$$\sum_i |\psi_i(t+\Delta t)|^2 = \sum_i |\psi_i(t)|^2. \tag{3-5}$$