CHAPTER 5. DEVICES AND MATRICES

1. Introduction

In the preceding chapters we have discussed the effect of various devices (analyzers, projectors, etc.) on particular beams of polarized photons. In the present chapter this discussion is generalized. We shall show that any device that acts only on the polarization state of a beam can be represented by a set of four complex amplitudes, conveniently displayed in a two-by-two array, or matrix. The four amplitudes are called matrix elements.

A single matrix describes a particular device in a particular representation; that is, it specifies directly the effect of the device on a beam of photons in one or the other of a particular set of basis states. If the matrix elements of a device are known in one representation, then a simple procedure of matrix multiplication provides the matrix elements of the same device in any other representation. Consequently, once the matrix of a device is given in any representation, the device is uniquely characterized. Moreover, the use of matrices makes it easy to determine the resultant effect of any sequential combination of devices whose individual matrices are known.

2. The matrix of a device

A device that affects only the polarization properties of a beam may be considered a "black box" with a single entrance hole through which an input beam enters, and a single exit hole through which an output beam emerges. The output beam has the same direction and energy as the input, but the intensity or the polarization state of the emerging beam may be different. Some devices that meet these criteria are:

- open hole
- beam stop (device that absorbs the entire beam)
- absorbing filter (device that absorbs some photons but does not otherwise change the characteristics of the transmitted beam)
- sheet polarizer (linear or circular polarizer)
- quarter wave plate
- half wave plate
- analyzer loop:
  - with both channels open,
  - or with just one channel open (projector),
  - or with neither channel open (beam stop),
  - or with an absorbing filter in one or both channels,
  - or with retarding plates in one or both channels.
- any sequential combination of the devices listed above.

Suppose we are given such a device as a "black box," and wish to investigate its properties.

One straightforward set of experiments that we can carry out is the following (Experiments 1 to 4): Let x-polarized beams and y-polarized beams in turn enter the input channel. In each case test the output beam first with an x-projector and then with a y-projector. One of the four experiments is sketched in Fig. 1, in which the box labeled A represents the device under study. For the other three experiments the beam stops are moved into different channels of the input and output analyzers.

Experiments 1 to 4 determine four real numbers (probabilities). This set of numbers is generally not sufficient to characterize the device uniquely. It is easy to exhibit two devices that give identical results in all four experiments, and yet have a different effect on some other incident beam. For example, suppose device A is an R-projector and device B is an L-projector. For both of these devices, each of the experiments 1 to 4 gives a probability 1/2. Yet the two devices are certainly distinct, as one can verify by passing an L-polarized beam through them.

The fact that experiments 1 to 4 are insufficient to characterize a device uniquely should not be surprising in view of the discussion of Chapter 4. That discussion suggests that we describe each measurement by a complex probability amplitude, of which only the magnitude has been determined in the present experiments. We shall see that four complex amplitudes indeed provide the desired unambiguous characterization of a device.

The notation used for these amplitudes is a generalization of the bracket notation introduced in the last chapter for projection amplitudes. The symbol for the amplitude must specify the initial state, the final state, and the device. We write it thus:

$$\langle \text{final state} \,|\, \text{device} \,|\, \text{initial state} \rangle \quad \text{or, more briefly} \quad \langle f \,|\, A \,|\, i \rangle \tag{1}$$

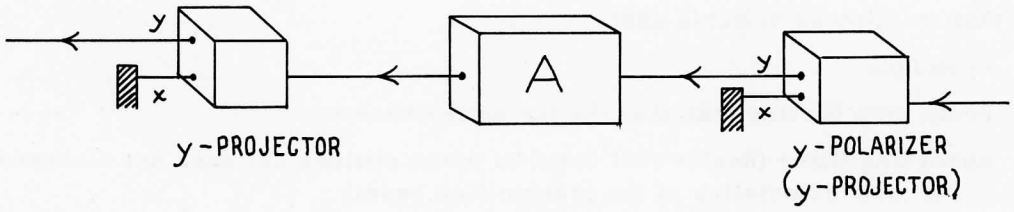


Figure 1. Experiment which measures the quantity $|\langle y \,|\, A \,|\, y \rangle|^2$. To measure $|\langle x \,|\, A \,|\, y \rangle|^2$, the stop is moved from the x channel of the final analyzer to the y channel.

The experiment of Figure 1 measures the absolute square of the amplitude $\langle y \,|\, A \,|\, y \rangle$. There are four such amplitudes, which are conventionally written in the form of a two-by-two array or matrix, as follows:

$$\begin{pmatrix} \langle x \,|\, A \,|\, x \rangle & \langle x \,|\, A \,|\, y \rangle \\ \langle y \,|\, A \,|\, x \rangle & \langle y \,|\, A \,|\, y \rangle \end{pmatrix} \tag{2}$$

Matrix (2) is the matrix of device A in the xy representation. The individual amplitudes that appear in the matrix are called matrix elements.

If four complex matrix elements are required in order to characterize a device, we must determine eight numbers (the real and imaginary parts of each amplitude, or alternatively the magnitude and phase of each one). This suggests that eight independent measurements are required in order to determine what is inside a "black box." However, it turns out that one phase can always be fixed arbitrarily. To see this, consider a general experiment of which device A constitutes one stage (Figure 2a). The boxes that precede and follow A in the diagram are arbitrary; each may contain alternative paths. The experiment remains essentially unchanged if open analyzer loops are inserted before and after A (Figure 2b), because open analyzer loops

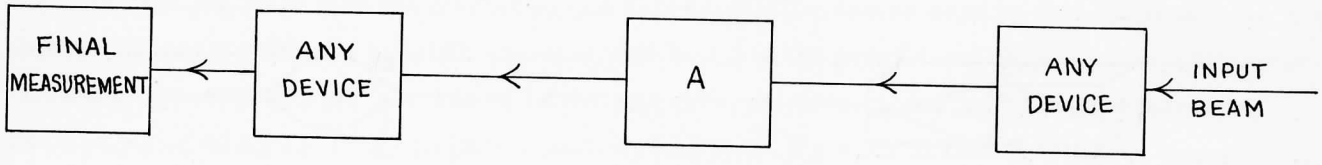


Figure 2a. General experiment used in the proof that the matrix elements of device A always have an overall phase undetermined by experiment.

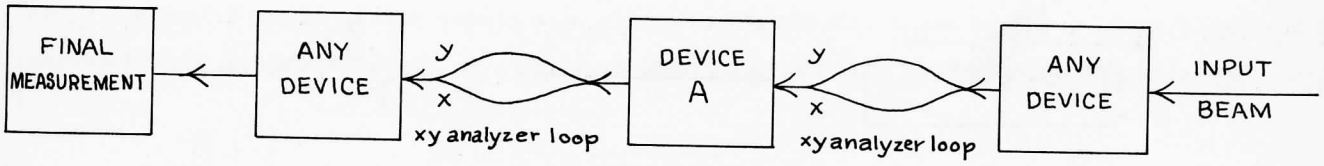
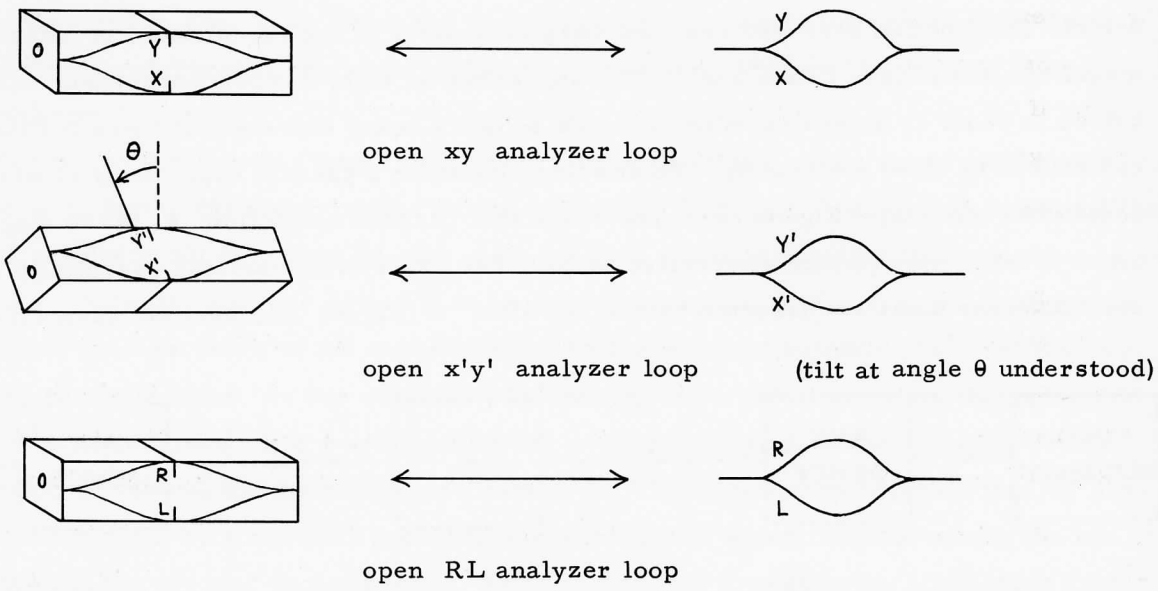


Figure 2b. A modified experiment that gives the same final measurement as the experiment of Figure 2a. Two xy analyzer loops have been inserted, one before A and one after A. New symbols for analyzer loops are explained in Figure 3.

transmit beams unchanged. The amplitude for this modified experiment can be written according to rules 1 and 2 of Chapter 4 (page 4-10). The amplitude consists of a sum of terms corresponding to the various possible paths through the complete system. In each possible path the passage through device A is represented by a matrix element of A in the xy representation. Hence each term in the sum contains as a factor exactly one matrix element of device A. Now suppose we were to replace the matrix of A by another matrix in which each element is $e^{i\alpha}$ times the corresponding element of the given matrix (α is a real number). Then each term in the sum that expresses the amplitude for the measurement in Figure 2 will be multiplied by the same phase factor $e^{i\alpha}$. When we take the absolute square of the resultant amplitude to obtain the probability that describes the experimental result, the common phase factor disappears. Hence the new matrix leads to exactly the same prediction as the old one for any experiment in which a device A is involved. It follows that two matrices related by a single over-all phase

Open analyzer loops

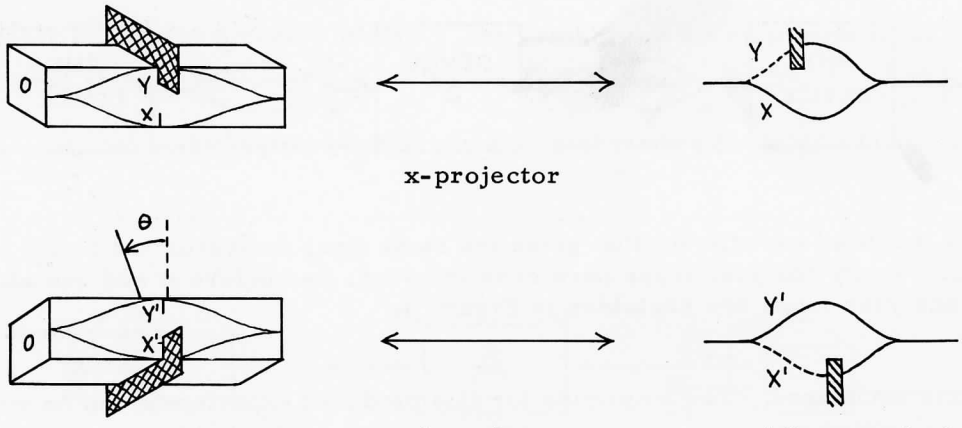


open xy analyzer loop

open x'y' analyzer loop (tilt at angle θ understood)

open RL analyzer loop

Projectors (examples)



x-projector

y'-projector (tilt at angle θ understood)

Figure 3. Simplified symbols for open analyzer loops and projectors to be used in this chapter and hereafter.

factor describe the same device, and furthermore that we cannot hope to determine the matrix elements of a given device except within a common phase factor. Thus, at most, seven independent measurements are required in order to identify a device. Generally, fewer than seven measurements suffice.

In Section 4 we shall show how the matrix elements can be evaluated by measuring the magnitudes of the matrix elements in two different representations. In order to carry out this program, we need to be able to relate the matrix elements in one representation to those in another representation. This change of basis is described in the next section.

3. Change of basis

How one determines the matrix elements of a device is discussed in Section 4. In the present section we show how the matrix in one representation can be used to find the matrix in any other representation. A specific example will indicate the procedure. Suppose the matrix elements of some device A are known in the xy representation, i.e., we know the values of $\langle x|A|x\rangle$, $\langle x|A|y\rangle$, $\langle y|A|x\rangle$, and $\langle y|A|y\rangle$. We wish to find the matrix elements of the same device in the RL representation. The experiment sketched in Figure 4a measures the magnitude of the element $\langle R|A|R\rangle$. At the present stage we cannot directly relate the outcome of this experiment to the known matrix elements in the xy representation. However, consider the modification of this experiment shown in Figure 4b. Open analyzer loops have been inserted before and after device A. Since an open analyzer loop by definition has no effect on any

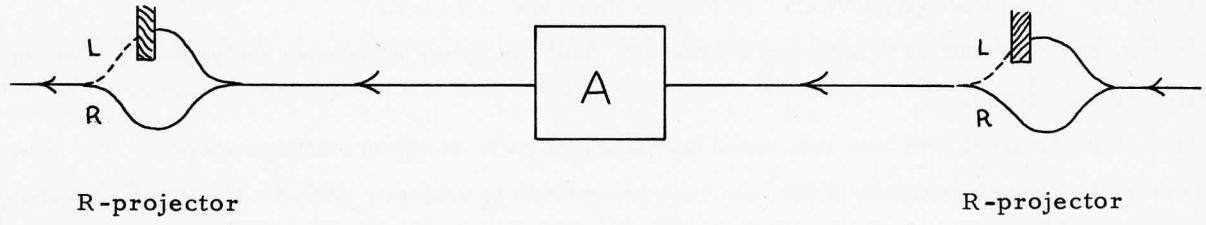


Figure 4a. Experiment which measures the quantity $|\langle R|A|R\rangle|^2$.

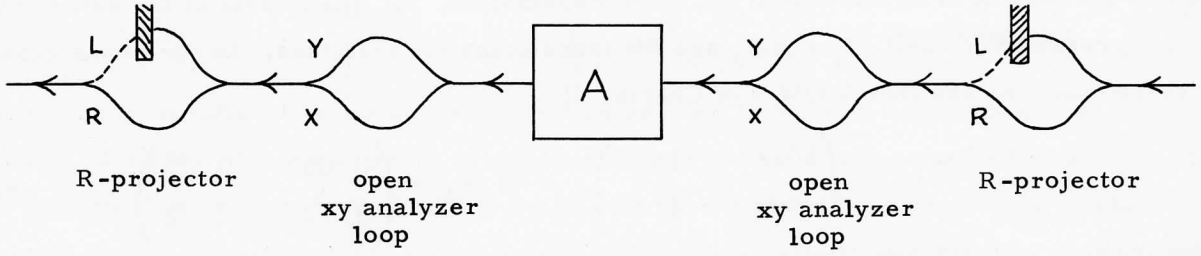


Figure 4b. Addition of open xy analyzer loops to experiment of Figure 4a.

beam, the result of the experiment of Figure 3b must be the same as the result of the experiment of Figure 3a. But we know how to analyze the experiment of Figure 3b by expanding the probability amplitude in terms of the amplitudes for all possible paths, using rules 1 and 2 of Chapter 4. There are four paths. The amplitude for proceeding, for example, by way of the x channel of the first analyzer loop and the y channel of the second loop is $\langle R|y\rangle\langle y|A|x\rangle\langle x|R\rangle$. The complete expression obtained by adding the comtributions of all four paths is

$$\begin{aligned}\langle R|A|R\rangle = {} & \langle R|x\rangle\langle x|A|x\rangle\langle x|R\rangle + \langle R|x\rangle\langle x|A|y\rangle\langle y|R\rangle \\ & + \langle R|y\rangle\langle y|A|x\rangle\langle x|R\rangle + \langle R|y\rangle\langle y|A|y\rangle\langle y|R\rangle\end{aligned} \tag{3}$$

Equation 3, together with the three similar expression for $\langle R|A|L\rangle$, $\langle L|A|R\rangle$ and $\langle L|A|L\rangle$, obtained in the same way, express the desired matrix elements in terms of the given ones and the projection amplitudes $\langle R|x\rangle$, $\langle x|L\rangle$, etc., whose values we know from Chapter 4. We have therefore verified the assertion that, given the matrix of a device in any representation, the effect of the device on a beam of photons in an arbitrary state may be determined.

The equations just derived, that express the transformation of a matrix from one representation to another, can be written compactly as a matrix multiplication*

$$\begin{pmatrix}\langle R|A|R\rangle & \langle R|A|L\rangle \\ \langle L|A|R\rangle & \langle L|A|L\rangle\end{pmatrix}$$

$$= \begin{pmatrix}\langle R|x\rangle & \langle R|y\rangle \\ \langle L|x\rangle & \langle L|y\rangle\end{pmatrix}\begin{pmatrix}\langle x|A|x\rangle & \langle x|A|y\rangle \\ \langle y|A|x\rangle & \langle y|A|y\rangle\end{pmatrix}\begin{pmatrix}\langle x|R\rangle & \langle x|L\rangle \\ \langle y|R\rangle & \langle y|L\rangle\end{pmatrix} \tag{4}$$

Equation 3 is one element of this matrix equation, and the other elements represent the other three similar equations.

Notice that the first and last matrices on the right side of (4) are independent of A; they characterize the transformation from one representation to another and are therefore called transformation matrices. The projection amplitudes that appear in these matrices are often called transformation coefficients. We may write Eq. 4 symbolically as

$$A' = U_1 A U_2 \tag{5}$$

where A is the matrix of any device in the xy representation, A' the matrix of the same device in the RL representation, and U_1 and U_2 are the transformation matrices. In our phase convention these matrices are (See Table 1 of Chapter 4)

$$U_1 = \begin{pmatrix}i/\sqrt{2} & 1/\sqrt{2} \\ -i/\sqrt{2} & 1/\sqrt{2}\end{pmatrix} \quad ; \quad U_2 = \begin{pmatrix}-i/\sqrt{2} & i/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix} \tag{6}$$

It may be verified by matrix multiplication that U_1 and U_2 are the inverses of one another; that

---

* See, for example, J. R. Munkres, Elementary Linear Algebra, Addison-Wesley Publishing Company, 1964.

is
$$U_1 U_2 = U_2 U_1 = I \tag{7}$$
where
$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{8}$$
is the identity or unit matrix. The relation (7), verified here in a specific example, is a general one. Hence we can rewrite Eq. 5 more simply as

$$A' = U\ A\ U^{-1} \tag{9}$$

where U is the matrix previously called U_1. U^{-1} is evidently the matrix that accomplishes the reverse transformation, from the RL representation to the xy representation. We can obtain the transformation from A' to A by multiplying both sides of (9) on the left by U^{-1} and on the right by U, and using (7) :

$$A = U^{-1} A'\ U \tag{10}$$

The same procedure enables us to transform a matrix from any representation to any other, as long as the corresponding projection amplitudes are known. For example, the matrix $U(xy \longrightarrow x'y')$ that transforms from the xy representation to the x'y' representation is

$$U(xy \longrightarrow x'y') = \begin{pmatrix} \langle x' | x \rangle & \langle x' | y \rangle \\ \langle y' | x \rangle & \langle y' | y \rangle \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{11}$$

Its inverse is

$$U(x'y' \longrightarrow xy) = U^{-1}(xy \longrightarrow x'y') = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{12}$$

Matrix (11) is exactly the same as the matrix that transforms the components of a vector from one set of rectangular coordinates to another set whose axes have been rotated through an angle θ. Equation 12 transforms these components back again. This was one reason for the particular choice of phase convention for projection amplitudes adopted in Chapter 4.

4. The matrix of a device

The transformation equations derived in the preceding section provide a general method that can be used to determine the matrix of any device. Suppose that experiments 1 through 4 of Section 2 have been carried out. We then know the magnitudes of the four matrix elements of device A in the xy representation, and can write the matrix in the form

$$A = \begin{pmatrix} a\, e^{i\alpha} & b\, e^{i\beta} \\ c\, e^{i\gamma} & d\, e^{i\delta} \end{pmatrix} \quad \text{xy representation} \tag{13}$$

where a, b, c, and d are known real numbers and α, β, γ, and δ are phases as yet undetermined. If all the elements in (13) are nonzero there are four unknowns, although according to the argument in Section 2 we expect to be able to fix the matrix elements only within a single overall phase. If one or more of the matrix elements vanishes, the number of unknowns is reduced accordingly.

Suppose we carry out four more experiments like 1 through 4, this time using some other

pair of basis states, e.g., R and L. The new set of measurements determines the magnitudes of four new matrix elements, viz. $\langle R|A|L\rangle$, etc. But Eq. 4 expresses these matrix elements in terms of the matrix elements in the xy representation, plus known transformation matrices. By taking the absolute squares of the four equations contained in this matrix equation, we obtain four equations in which the only unknowns are the desired phases. There must be at least one redundant equation in the set; otherwise we could determine all phases uniquely. If the equations have a unique solution for the relative phases, our task is accomplished. Sometimes there is not a unique solution (See example 2c below). In that case a different representation must be used.

If all the elements in the original matrix (13) are non-zero, the algebraic solution for the phases may be rather complicated. For example, the equation one obtains by substituting (6) and (13) into (3) and squaring is

$$\left|\langle R|A|R\rangle\right|^2 = \frac{1}{4}\left|ae^{i\alpha} + ibe^{i\beta} - ice^{i\gamma} + de^{i\delta}\right|^2 \qquad (14)$$

When expanded, Eq. 14 contains sines and cosines of the various relative phases $\alpha - \beta$, $\alpha - \gamma$, etc. However, for most of the devices with which we are concerned, there are sufficient zero matrix elements to make the determination of phases quite simple. A few examples illustrate the procedure.

Example 1. Projectors

If the device in question is an x-projector, the results of experiments 1 through 4 are very simple. Three of the measurements yield the result zero, and the fourth gives probability unity (Figure 5). Letting P_x denote the x-projector, we can write

$$\left|\langle x|P_x|x\rangle\right|^2 = 1;\ \left|\langle x|P_x|y\rangle\right|^2 = \left|\langle y|P_x|x\rangle\right|^2 = \left|\langle y|P_x|y\rangle\right|^2 = 0 \qquad (15)$$

For this device there are no unknown phases at all to be discussed. The only quantity unspecified is the phase of the single nonzero matrix element $\langle x|P_x|x\rangle$. But this constitutes only an over-all phase for the entire matrix which, as we have already pointed out, cannot be determined. Therefore we may fix this phase at our pleasure. The simplest choice is to put the matrix element equal to unity. Thus

$$P_x = \begin{pmatrix} \langle x|P_x|x\rangle & \langle x|P_x|y\rangle \\ \langle y|P_x|x\rangle & \langle y|P_x|y\rangle \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad (16)$$

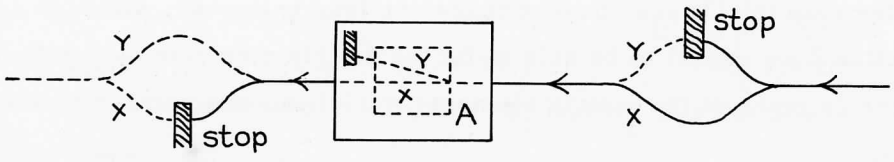


Figure 5. Experiment to measure $\left|\langle y|A|x\rangle\right|^2$ for an x-projector. For this experiment there is zero output.

This is a rather trivial-appearing result. However, using the results of the preceding sections, we can immediately express P_x in any other representation, in which the matrix has a far less trivial appearance. For example, using Eqs. 5 and 6, we find

$$P_x = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix} \qquad \text{RL representation} \tag{17}$$

and similarly

$$P_x = \begin{pmatrix} \cos^2\theta & -(\cos\theta \sin\theta) \\ -(\cos\theta \sin\theta) & \sin^2\theta \end{pmatrix} \qquad \text{x'y' representation} \tag{18}$$

From the definition of a projector, it follows that the matrix of <u>any</u> projector is given by Eq. 16 in the representation in which the state defined by the projector is one of the basis states. Thus for an R-projector, denoted by P_R, we can write immediately

$$P_R = \begin{pmatrix} \langle R|P_R|R\rangle & \langle R|P_R|L\rangle \\ \langle L|P_R|R\rangle & \langle L|P_R|L\rangle \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{RL representation} \tag{19}$$

and, using Eq. 5,

$$P_R = \begin{pmatrix} 1/2 & -i/2 \\ i/2 & 1/2 \end{pmatrix} \qquad \text{xy representation} \tag{20}$$

<u>Example 2.</u>

The next simplest case is one in which two matrix elements are zero. Suppose that for some device A, experiments 1 to 4 of Section 2 give the following results:

$$|\langle x|A|x\rangle|^2 = |\langle y|A|y\rangle|^2 = 1; \qquad |\langle x|A|y\rangle|^2 = |\langle y|A|x\rangle|^2 = 0 \tag{21}$$

That is, device A transmits both state x and state y with undiminished intensity. The device <u>might</u> be an open hole; but, as we shall see, it need not be. From the results (21), we write the general matrix (13) as

$$A = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{i\delta} \end{pmatrix} \qquad \text{xy representation} \tag{22}$$

There is only one relative phase to be determined. To do so, we measure the magnitudes of the matrix elements in the RL representation by appropriate modification of experiments 1 through 4. Equation 14 with b and c set equal to zero reads

$$|\langle R|A|R\rangle|^2 = (1/2)\,[1 + \cos(\alpha - \delta)] \tag{23}$$

and one can obtain three similar equations for $\langle R|A|L\rangle$, etc., all of which involve the phase difference $\alpha - \delta$. We consider three alternative results.

<u>Case 2a:</u> $|\langle R|A|R\rangle|^2 = 1$. The device transmits an R-polarized beam with undiminished intensity. Equation 23 tells us that $\cos(\alpha - \delta) = 1$, which implies $e^{i\alpha} = e^{i\delta}$. We can put this common phase factor equal to unity. Then the device in question has the matrix

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{xy representation} \tag{24}$$

which is the unit matrix. The device must be an open hole or an open analyzer loop or some such similar device. The experiments we have performed (notice they are five in number) are sufficient to determine that the device must transmit any other incident beam unchanged. When the unit matrix is transformed to any other representation, its form remains unchanged. No other matrix has this property.

Case 2b. $|\langle R|A|R\rangle|^2 = 0.$ In this case we find from Eq. 23 that $\cos(\alpha - \delta) = -1$, which implies that $e^{i\alpha} = -e^{i\delta}$. The matrix can be written

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{xy representation} \tag{25}$$

One device that has this matrix is a half-wave plate with a particular orientation (see the exercises).

Case 2c. Quarter wave plate. Suppose the device A happens to be a quarter wave plate with its fast axis oriented along the x direction. (See Chapter 3, Box 2 for definition and properties of this device.) From the properties of the quarter wave plate we can predict, and experiment readily verifies, that experiments 1 through 4 in the xy representation yield just the results (21). Thus the matrix of A must be of the form (22). Moreover, we can also predict that the second set of experiments, with the RL representation, yield the following results:

$$|\langle R|A|R\rangle|^2 = |\langle R|A|L\rangle|^2 = |\langle L|A|R\rangle|^2 = |\langle L|A|L\rangle|^2 = \frac{1}{2} \tag{26}$$

These predictions are likewise verifiable by experiment. When we substitute these values in Eq. (23) and the three other similar equations, we find that

$$\cos(\alpha - \delta) = 0 \tag{27}$$

which leads to two distinct solutions for the phase factors:

$$e^{i\delta} = ie^{i\alpha} \tag{28a}$$

and

$$e^{i\delta} = -ie^{i\alpha} \tag{28b}$$

No further phase information can be gleaned from these experiments; we therefore cannot decide, on the basis of these data, whether the matrix of A ought to be writte

$$\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \tag{29a}$$

or

$$\begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix} \tag{29b}$$

In point of fact, (29b) is the correct matrix of A; (29a) represents a quarter wave plate with its fast axis along the y direction, which is a different device even though it gives the same results, (23) and (26), in both sets of experiments we have considered. An additional measurement is required in order to determine the correct matrices for these devices. (See the exercises.)

The example of a quarter wave plate illustrates that measurement of the magnitudes of the

matrix elements in any two representations will not, in general, fix the matrix of device (even within an over-all phase). However, one can always find two particular representations that lead to a unique determination. As already mentioned, the algebra can get quite complicated if an inconvenient choice is made.

It is possible--and is sometimes useful--to present a matrix in a mixed representation, that is, a form in which the input state is in a different representation than the output state. For example, the matrix element $\langle R|P_x|x\rangle$ for an x-projector is in a mixed representation of linear and circular polarization states. Calculation of a matrix in a mixed representation can be done by applying only one transformation matrix (Eq. 9) either the one on the left or the one on the right, depending on the mixture desired. For example, to calculate $\langle R|P_x|x\rangle$ we have

$$\langle R|P_x|x\rangle = \langle R|x\rangle\langle x|P_x|x\rangle + \langle R|y\rangle\langle y|P_x|x\rangle = i/\sqrt{2} \qquad (30)$$

It is important to notice that not every two-by-two matrix corresponds to a physical device. If we confine ourselves to passive devices, that is, ones that do not manufacture photons, then at least one limitation is clear: for any input beam, the sum of the intensities of the two states in the output beam is no greater than the intensity of the input. This means that the matrix elements must satisfy the following inequality

$$\sum_j |\langle j|A|i\rangle|^2 \leq 1 \qquad (31)$$

Here i denotes any state whatever, and the sum runs over the pair of basis states in any representation. Equation 31 implies that the sum of the absolute squares of the elements in any column of a matrix is at most unity. It is also true that the absolute squares of the elements in any row can also sum to no more than unity : $\sum_i |\langle j|A|i\rangle|^2 \leq 1$ (32)

It is worth while remarking that matrices can be used also in a completely classical description of the devices we have been discussing. An explanation of the classical meaning of these matrices--called Jones matrices--and a table of matrices for different devices (including those treated as examples here) appears in the book by Shurcliff and Ballard.*

5. The matrix for a series of consecutive devices

Let a beam pass sequentially through devices A and B (Figure 6). Suppose that matrix ele-

* Shurcliff and Ballard, op. cit. (reference on page 2-1), pages 80, 89, and 94.

ments for the two separate devices are known in, say, the xy representation. There is a simple way to summarize the net result of these two devices in series. In particular, we can find a matrix for the combination, treated as a single "device C" (dashed line, Figure 6). Symbolically, add an open xy analyzer loop between A and B, as shown in Figure 7. There are two alternative paths through this system. The resultant amplitude is simply $\langle x|C|x\rangle$:

$$\langle x|C|x\rangle = \langle x|B|x\rangle \langle x|A|x\rangle + \langle x|B|y\rangle \langle y|A|x\rangle \tag{33}$$

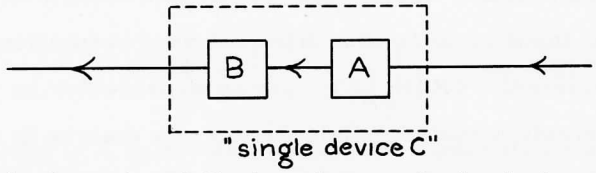


Figure 6. The sequential devices A and B viewed as a single device C (dashed line).

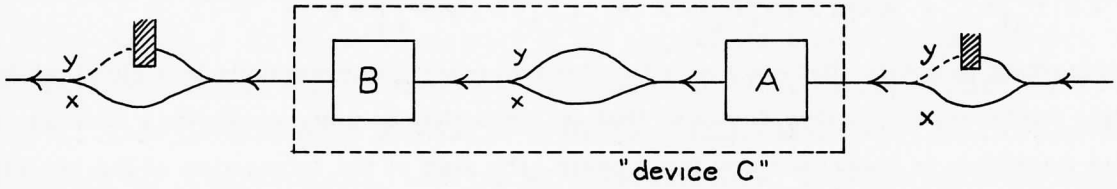


Figure 7. Thought experiment useful in finding the matrix elements of C from the matrix elements of A and B.

All amplitudes on the right side of this equation are assumed known. Therefore the matrix element $\langle x|C|x\rangle$ for the single device C is now known. The remaining three matrix elements are found from experiments similar to that of Figure 7. These four equations for the matrix elements of C can be summarized in terms of the matrix multiplication of matrices B and A:

$$\begin{pmatrix} \langle x|C|x\rangle & \langle x|C|y\rangle \\ \langle y|C|x\rangle & \langle y|C|y\rangle \end{pmatrix} = \begin{pmatrix} \langle x|B|x\rangle & \langle x|B|y\rangle \\ \langle y|B|x\rangle & \langle y|B|y\rangle \end{pmatrix} \begin{pmatrix} \langle x|A|x\rangle & \langle x|A|y\rangle \\ \langle y|A|x\rangle & \langle y|A|y\rangle \end{pmatrix} \tag{34}$$

or, more simply

$$C = B\,A \tag{35}$$

The last equation is true for any single representation. Since matrices do not always commute (AB not necessarily equal to BA), one must be careful to write the matrices in the same order as the devices. The beam direction for most of the figures in this book has been chosen from right to left, so the order of devices is the same as the order of multiplication of matrices for the devices.

Three or more devices in series are easily described as a single device using a simple extension of the method of matrix multiplication. The matrix for the combined device is the product of all the individual matrices in the correct order.

As an example, consider the combination of devices shown in Figure 8. Using the matrices

of the individual devices derived in Section 4, we find that the matrix for the combination of devices operating in the order shown in the figure is (in the xy representation)

$$BA = \begin{pmatrix} 1/2 & -i/2 \\ i/2 & 1/2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 \\ i/2 & 0 \end{pmatrix} \tag{36}$$

If the order of the devices is reversed, the matrix of the combination is

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/2 & -i/2 \\ i/2 & 1/2 \end{pmatrix} = \begin{pmatrix} 1/2 & -i/2 \\ 0 & 0 \end{pmatrix} \tag{37}$$

other examples are presented in the exercises.

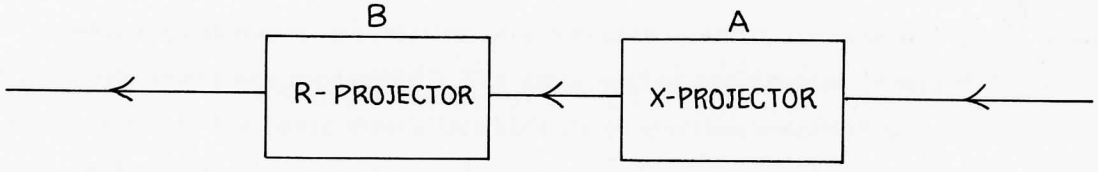


Figure 8. A series of devices: an example of the series shown in Figure 6.

CHAPTER 6. STATE VECTORS AND OPERATORS

1. Introduction

This chapter completes our presentation of the formal apparatus of quantum mechanics. No new experiments are considered in this chapter, but the experiments of preceding chapters are reinterpreted in terms of two powerful concepts: state vector and operator. The state vector concept is not only useful for describing photon states but also will be immediately applicable to atomic beam states, as described in Chapter 7.*

A state vector receives its name from the strong analogy between its properties and those of an ordinary unit vector. The components of an ordinary unit vector are equal to its projections on a given set of mutually perpendicular axes. The components of a state vector are its projection amplitudes onto a given set of mutually orthogonal basis states. The ordinary unit vector provides an abstract description of a directed quantity independent of the axes along which its components are measured. The state vector provides an abstract description of a state independent of the basis states in which its projection amplitudes are measured.

An operator provides an abstract description of a device. An operator changes one state vector into another. The matrices introduced in the preceding chapter constitute a concrete representation of operators. In a similar manner row and column matrices constitute a concrete representation of the abstract state vectors.

2. The state vector

For a heuristic introduction to the state vector, we utilize once more the familiar thought-experiment of Figure 1. Let ψ denote any state of photon polarization (that is, ψ can stand for y or y' or R, etc.). With the help of the open analyzer loop, we can express the projection amplitude from the state ψ to the state of linear polarization y' as:

$$\langle y' | \psi \rangle = \langle y' | y \rangle \langle y | \psi \rangle + \langle y' | x \rangle \langle x | \psi \rangle \qquad (1a)$$

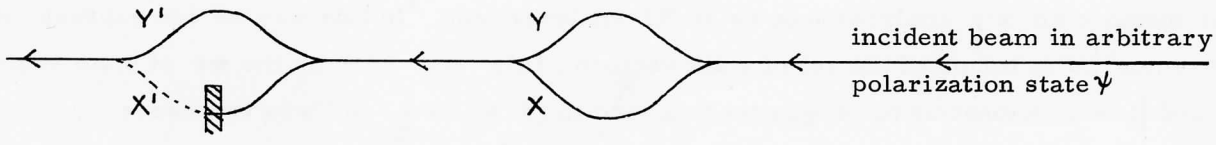


Figure 1. Experiment used to introduce the state vector.

* In contrast the operator concept will not be used again for several chapters, so may be omitted from a first reading of this chapter if the instructor so desires.

Alternatively, by replacing the final projector in Figure 1 with a differently oriented linear projector or with an R-projector we obtain similar equations

$$\langle y''|\psi\rangle = \langle y''|y\rangle\langle y|\psi\rangle + \langle y''|x\rangle\langle x|\psi\rangle \quad (1b)$$

$$\langle R|\psi\rangle = \langle R|y\rangle\langle y|\psi\rangle + \langle R|x\rangle\langle x|\psi\rangle \quad (1c)$$

The only difference between the three experiments above is the identity of the final projector, The only difference in the corresponding equations is the symbol that represents the final projector in the appropriate amplitudes. We can summarize all three equations (as well as all other equations of the same type)by writing

$$\langle\ |\psi\rangle = \langle\ |y\rangle\langle y|\psi\rangle + \langle\ |x\rangle\langle x|\psi\rangle \quad (2)$$

The blank spaces in Eq. 2 may be filled in by <u>any</u> symbol that represents a photon polarization state (the same symbol goes in all the blanks). The resulting equation then specifies the projection amplitude from the state ψ to the state in question, expressed in terms of the projection amplitudes from ψ to the x and y basis states. The notation can be abbreviated by removing the angular braces which enclose the blanks; thus instead of (2) we write

$$|\psi\rangle = |y\rangle\langle y|\psi\rangle + |x\rangle\langle x|\psi\rangle \quad (3)$$

For example, if ψ stands for the state R, Eq. 3 reads

$$|R\rangle = |x\rangle\langle x|R\rangle + |y\rangle\langle y|R\rangle = |x\rangle(-i/\sqrt{2}) + |y\rangle(1/\sqrt{2}) \quad (4)$$

The symbol $|\psi\rangle$ is called the <u>state vector</u> for the state ψ. The analogy to ordinary vectors is elaborated below. Because the $|\ \rangle$ is the right half of the Dirac bracket $\langle\ |\ \rangle$, it is sometimes called a <u>ket</u> or <u>ket vector.</u>

One may of course regard Eq. 3 as merely a shorthand form of (2), which in turn represents (1a), (1b), (1c) and all such equations. However, a more significant meaning can be attached to Eq. 2. This equation corresponds to the right-hand half of Figure 1: a beam in state ψ and an analyzer loop. The analyzer loop has no net effect on the beam, merely analyzing the beam into the x, y basis states and then recombining them again. In some sense Eq. 3 therefore stands for the state ψ. We regard the state vector $|\psi\rangle$ as an abstract symbol for the state ψ. The right side of (3) expresses this state vector in terms of the state vectors $|x\rangle$ and $|y\rangle$ that symbolize the basis states x and y. We can just as well use some other analyzer loop in Figure 1, for instance an x'y' analyzer loop or an RL analyzer loop. In this way we can express the state vector $|\psi\rangle$ in terms of the set of state vectors $|x'\rangle$ and $|y'\rangle$ or the set of state vectors $|R\rangle$ and $|L\rangle$. Repeating the argument that lead to (3) we have, in these two cases

$$|\psi\rangle = |x'\rangle\langle x'|\psi\rangle + |y'\rangle\langle y'|\psi\rangle \quad (5)$$

$$|\psi\rangle = |R\rangle\langle R|\psi\rangle + |L\rangle\langle L|\psi\rangle \quad (6)$$

The expressions (3), (5), and (6) are all analogous to the expansion of an ordinary vector in terms of unit vectors that point along mutually perpendicular axes. The analogy is most complete if we limit consideration to <u>unit vectors in a plane.</u> Such a unit vector $\hat{A}$ can be written

$$\hat{A} = \hat{x} A_x + \hat{y} A_y \tag{7}$$

where $\hat{x}$ and $\hat{y}$ denote unit vectors in the x and y directions, and A_x and A_y are the components of $\hat{A}$ along these axes. Alternatively one can expand the same unit vector $\hat{A}$ in terms of some other pair of perpendicular unit vectors x' and y':

$$\hat{A} = \hat{x}' A_{x'} + \hat{y}' A_{y'} \tag{8}$$

The right sides of (7) and (8) represent the same vector in two different explicit ways. In contrast the symbol $\hat{A}$ on the left sides of (7) and (8) is an abstract way of describing the vector without reference to any particular set of axes.

The symbol $|\psi\rangle$ is an abstract symbol for the quantum state in the same way as the symbol $\hat{A}$ is an abstract symbol for a unit vector. The state vector $|\psi\rangle$ is independent of any particular representation. The "components" of $|\psi\rangle$ in a particular representation are, as Eqs. 3, 5, and 6 indicate, the projection amplitudes from the state ψ to the basis states of the representation. The projection amplitudes to any complete set of basis states uniquely define the state (once the phase convention has been fixed) just as the components of a vector $\hat{A}$ with respect to a given set of axes uniquely define the vector.

Notice that the projection amplitudes from any state ψ to the basis states x and y satisfy the relation ("completeness:" page 3-27):

$$|\langle x | \psi \rangle|^2 + |\langle y | \psi \rangle|^2 = 1 \tag{9}$$

just as the components of any unit vector $\hat{A}$ along any pair of axes satisfy the relation ("unit vector"):

$$A_x^2 + A_y^2 = 1 \tag{10}$$

But, of course, the "components" of a state vector are in general complex.

The analogy between state vectors and unit vectors extends even further. If the components of a unit vector along a given set of axes are known, then the components along any other set of axes can be determined. Specifically, if the x'y' axes make an angle θ with the xy axes (Figure 2), then the components of A along the primed axes are:

$$\begin{aligned} A_{x'} &= \cos\theta\, A_x + \sin\theta\, A_y \\ A_{y'} &= -\sin\theta\, A_x + \cos\theta\, A_y \end{aligned} \tag{11}$$

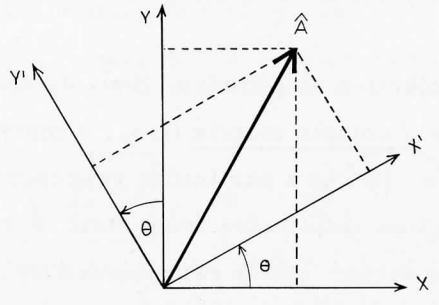


Figure 2. Relations between components of $\hat{A}$ in primed and unprimed coordinate systems.

The transformation law (11) is an essential property of vectors: In fact, it can even be taken as the definition of an ordinary vector: an ordinary vector is a quantity with two components that transform under a rotation of axes according to Eq. (11). The statement that some physical quantity like force or momentum is a vector means that its components obey the transformation law (11). It follows that if a physical law is expressed as a relation between vector quantities, the law is automatically invariant under rotations. (All the terms transform in exactly the same way.)

The components of a state vector likewise obey well-defined transformation laws. From simple experiments with analyzer loops we can, for example, derive the following transformation laws:

a) transformation from one set of linearly polarized basis states to another such set. We have already seen that

$$\begin{aligned} \langle x'|\psi\rangle &= \langle x'|x\rangle\langle x|\psi\rangle + \langle x'|y\rangle\langle y|\psi\rangle \\ \langle y'|\psi\rangle &= \langle y'|x\rangle\langle x|\psi\rangle + \langle y'|y\rangle\langle y|\psi\rangle \end{aligned}$$

Therefore, using Table 1 of Chapter 4 we have

$$\begin{aligned} \langle x'|\psi\rangle &= \cos\theta\,\langle x|\psi\rangle + \sin\theta\,\langle y|\psi\rangle \\ \langle y'|\psi\rangle &= -\sin\theta\,\langle x|\psi\rangle + \cos\theta\,\langle y|\psi\rangle \end{aligned} \tag{12}$$

b) transformation from linearly polarized basis states to circularly polarized basis states:

We know

$$\begin{aligned} \langle R|\psi\rangle &= \langle R|x\rangle\langle x|\psi\rangle + \langle R|y\rangle\langle y|\psi\rangle \\ \langle L|\psi\rangle &= \langle L|x\rangle\langle x|\psi\rangle + \langle L|y\rangle\langle y|\psi\rangle \end{aligned}$$

Thus

$$\begin{aligned} \langle R|\psi\rangle &= (i/\sqrt{2})\langle x|\psi\rangle + (1/\sqrt{2})\langle y|\psi\rangle \\ \langle L|\psi\rangle &= (-i/\sqrt{2})\langle x|\psi\rangle + (1/\sqrt{2})\langle y|\psi\rangle \end{aligned} \tag{13}$$

The transformation law (12) is precisely the same as (11): The components of a photon state vector with respect to linear polarization basis states transform under a rotation of axes exactly as do the components of a vector in 2-space. This identity holds only with the particular phase convention we have adopted. (It is one of the reasons for choosing this convention.) But there is no transformation of vectors that corresponds to the transformation (13) from linear to circular polarization. In this respect state vectors are more general than ordinary unit vectors. The important point is that a well-defined transformation law exists for any pair of representations. Because of this, any equation relating state vectors is invariant to a change of basis.

The transformation laws for projection amplitudes, Eqs. 12 and 13, can be expressed conveniently in matrix form. We define a column matrix (i.e., a matrix with a single column and two rows) to describe the state vector $|\psi\rangle$ in a particular representation. The elements of this column matrix are the pair of projection amplitudes from state ψ to the basis states of the representation. For example the state vector $|\psi\rangle$ is represented by any of the column matrices

$$\begin{pmatrix} \langle x|\psi\rangle \\ \langle y|\psi\rangle \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \langle x'|\psi\rangle \\ \langle y'|\psi\rangle \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \langle R|\psi\rangle \\ \langle L|\psi\rangle \end{pmatrix}$$

In terms of these column matrices, the transformation equations (12) and (13) take the form of matrix multiplications

$$\begin{pmatrix} \langle x'|\psi\rangle \\ \langle y'|\psi\rangle \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \langle x|\psi\rangle \\ \langle y|\psi\rangle \end{pmatrix} \tag{14}$$

$$\begin{pmatrix} \langle R|\psi\rangle \\ \langle L|\psi\rangle \end{pmatrix} = \begin{pmatrix} i/\sqrt{2} & 1/\sqrt{2} \\ -i/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} \langle x|\psi\rangle \\ \langle y|\psi\rangle \end{pmatrix} \tag{15}$$

The two-by-two matrix in Eq. 14 is precisely the same as the transformation matrix $U(xy \rightarrow x'y')$ used in Chapter 5 to describe the transformation of the matrix of a device from one representation to another (Eq. 11 of Chapter 5). Similarly, the two-by-two matrix in (15) is $U(xy \rightarrow RL)$ given in Eq. 6 of Chapter 5. Notice the difference in form between transformation equations in this chapter and transformation equations in Chapter 5. In transforming the components of a state vector only a single transformation matrix U appears. In transforming a matrix, the matrix to be transformed is "sandwiched" between the transformation matrices U and U^{-1}.

A final streamlining of the notation is possible with the help of the summation symbol. We can write a single equation which expresses a state vector in terms of an arbitrary set of basis states:

$$|\psi\rangle = \sum_j |j\rangle\langle j|\psi\rangle \tag{16}$$

The dummy index j in the summation of Eq. 16 represents, in turn, the two symbols that describe the orthogonal basis states in any representation: e.g., x and y (this gives Eq. 3), or R and L (this gives Eq. 6), or x' and y' (this gives Eq. 5).

3. The dual state vector

Suppose we "turn around" the experiment in Figure 1, i.e., start with a beam which is y' polarized, pass it through an open xy analyzer loop and then through a "ψ-projector." (Figure 2.) (We saw in Chapter 3 that a projector must exist for any state.) This thought-experiment enables us to express the amplitude $\langle\psi|y'\rangle$ as

$$\langle\psi|y'\rangle = \langle\psi|y\rangle\langle y|y'\rangle + \langle\psi|x\rangle\langle x|y'\rangle \tag{17}$$

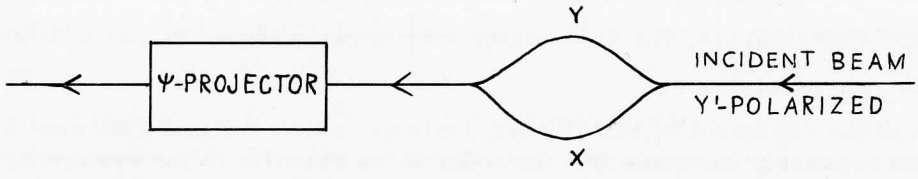


Figure 3. Thought-experiment used to introduce the dual state vector.

We can now repeat the argument of Section 2, imagining other beams incident on the same apparatus, and write a symbolic equation that summarizes all such similar experiments

$$\langle \psi | \; \rangle \;\; = \langle \psi | y \rangle \langle y | \; \rangle + \langle \psi | x \rangle \langle x | \; \rangle \qquad (18)$$

As in Eq. 2, the blank spaces are to be filled by the symbol representing the incident beam. Equation 18 can be abbreviated in the form

$$\langle \psi | \;\; = \;\; \langle \psi | y \rangle \langle y | \;\;\; + \langle \psi | x \rangle \langle x | \qquad (19)$$

The entire discussion of state vectors in Section 2 applies equally well to the symbol $\langle \; |$ introduced in Eq. 19. $\langle \psi |$ represents the state of ψ polarization and has properties similar to those of a state vector. The symbol $\langle \psi |$ is called the dual state vector, or more simply the dual vector. The word dual comes from the mathematical theory of vector spaces.* Because it forms the front part of the Dirac bracket, the dual state vector is sometimes called the bra vector.

The components of the dual vector $\langle \psi |$ in the xy representation are the probability amplitudes found in Eq. 19: $\langle \psi | y \rangle$ and $\langle \psi | x \rangle$. These are not the same as $\langle y | \psi \rangle$ and $\langle x | \psi \rangle$ respectively, the components of the vector $| \psi \rangle$ that describes the same state. Rather, the components of the state vector and the dual vector for the same state are complex conjugates of one another (Eq. 24 of Chapter 4). Likewise the transformation laws for components of a dual vector are related to, but not identical with, the transformation laws for components of a state vector. Instead of Eqs. 12 and 13 one obtains by the same procedure, for transformation to the x'y' basis:

$$\langle \psi | x' \rangle \;\; = \langle \psi | x \rangle \langle x | x' \rangle + \langle \psi | y \rangle \langle y | x' \rangle$$
$$\langle \psi | y' \rangle \;\; = \langle \psi | x \rangle \langle x | y' \rangle + \langle \psi | y \rangle \langle y | y' \rangle$$

therefore (Table 1 of Chapter 4)

$$\langle \psi | x' \rangle \;\; = \langle \psi | x \rangle \cos\theta + \langle \psi | y \rangle \sin\theta$$
$$\langle \psi | y' \rangle \;\; = \langle \psi | x \rangle (-\sin\theta) + \langle \psi | y \rangle \cos\theta \qquad (20)$$

and for transformation to the RL basis

$$\langle \psi | R \rangle \;\; = \langle \psi | x \rangle \langle x | R \rangle + \langle \psi | y \rangle \langle y | R \rangle$$
$$\langle \psi | L \rangle \;\; = \langle \psi | x \rangle \langle x | L \rangle + \langle \psi | y \rangle \langle y | L \rangle$$

therefore

$$\langle \psi | R \rangle \;\; = \langle \psi | x \rangle (-i/\sqrt{2}) + \langle \psi | y \rangle (1/\sqrt{2})$$
$$\langle \psi | L \rangle \;\; = \langle \psi | x \rangle (i/\sqrt{2}) + \langle \psi | y \rangle (1/\sqrt{2}) \qquad (21)$$

These equations can, like (12) and (13), be written conveniently in matrix form. However, in this case the dual vector must be written as a row matrix (one row, two columns) rather than a column matrix. For example, the dual vector $\langle \psi |$ is represented by the following row matrices in different representations:

* Actually each of the vectors $\langle \psi |$ and $| \psi \rangle$ are duals of each other and both are state vectors, since they both represent the state ψ. In order to be specific in the present treatment, we usually call the bra vector $\langle \psi |$ the dual vector, leaving the label vector or state vector for the ket $| \psi \rangle$.

$$\left(\langle \psi | x \rangle \quad \langle \psi | y \rangle \right) \quad \text{or} \left(\langle \psi | x' \rangle \quad \langle \psi | y' \rangle \right) \text{ or } \quad \left(\langle \psi | R \rangle \quad \langle \psi | L \rangle \right)$$

In terms of these row matrices, Eqs. 20 and 21 are represented by the matrix equations

$$\left(\langle \psi | x' \rangle \quad \langle \psi | y' \rangle \right) = \left(\langle \psi | x \rangle \quad \langle \psi | y \rangle \right) \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{22}$$

$$\left(\langle \psi | R \rangle \quad \langle \psi | L \rangle \right) = \left(\langle \psi | x \rangle \quad \langle \psi | y \rangle \right) \begin{pmatrix} -i/\sqrt{2} & i/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \tag{23}$$

Notice that according to the laws of matrix multiplication, a two-by-two matrix can only post-multiply a one-by-two row matrix; that is, the order of the matrices must be as it appears in (22) and (23). The transformation matrices in these expressions are the adjoints* of the U matrices of Eqs. 14 and 15. Thus eqs. (22) and (23) can be written as

$$\left(\langle \psi | x' \rangle \quad \langle \psi | y' \rangle \right) = \left(\langle \psi | x \rangle \quad \langle \psi | y \rangle \right) U^{\dagger}(xy \longrightarrow x'y') \tag{24}$$

$$\left(\langle \psi | R \rangle \quad \langle \psi | L \rangle \right) = \left(\langle \psi | x \rangle \quad \langle \psi | y \rangle \right) U^{\dagger}(xy \longrightarrow RL) \tag{25}$$

Analogously to Eq. 16 we can summarize Eqs. 20 and 21 and all similar equations using the summation symbol

$$\langle \psi | = \sum_j \langle \psi | j \rangle \langle j | \tag{26}$$

Here, as before, the dummy index j represents, in turn, the two symbols that describe the orthogonal basis states in any representation.

4. Inner products

Equation 3 expresses the state vector $|\psi\rangle$ in terms of the xy basis: $|x\rangle$ and $|y\rangle$. Equation 19 performs a similar task for the dual vector $\langle\psi|$. In both of these equations there are coefficients that are numbers. State vectors, as we have already emphasized, are not numbers but entirely different mathematical entities. An ordinary vector is also a mathematical entity different from a number. In vector algebra, however, there is a way of obtaining a number from two vectors by defining their inner or scalar product. The scalar product is usually called the dot product and written $\vec{A} \cdot \vec{B}$. For unit vectors $\hat{A}$ and $\hat{B}$ in two dimensions we have

$$\hat{A} \cdot \hat{B} = A_x B_x + A_y B_y \tag{27}$$

The scalar product is useful because it is invariant under rotation, that is

$$A_x B_x + A_y B_y = A_{x'} B_{x'} + A_{y'} B_{y'} \tag{28}$$

as can be verified directly using the transformation law (11). Moreover, the coefficients in the

---

* The adjoint of a matrix A, written $A^{\dagger}$, is obtained by interchanging the rows and columns and taking the complex conjugate of all the elements; that is, $A^{\dagger}_{ij} = (A_{ji})^*$. It happens that for the matrices involved here, the adjoint is identical to the inverse. A matrix with this property is called unitary. The transformations we are discussing are all unitary transformations.

expansion of a vector in a given coordinate system, i.e., the components of a vector along a particular pair of axes (Eqs. 7 and 8), may be expressed as the scalar products of the vector with unit vectors that point along the coordinate axes. For example

$$\begin{aligned} A_x &= \hat{x} \cdot \hat{A} = \hat{A} \cdot \hat{x} \\ A_y &= \hat{y} \cdot \hat{A} = \hat{A} \cdot \hat{y} \\ A_{x'} &= \hat{x}' \cdot \hat{A} = \hat{A} \cdot \hat{x}' \\ A_{y'} &= \hat{y}' \cdot \hat{A} = \hat{A} \cdot \hat{y}' \end{aligned} \tag{29}$$

(According to the definition (27) the order of factors in a scalar product of ordinary vectors is immaterial.) Thus Eq. 27 may be written:

$$\hat{A} \cdot \hat{B} = (\hat{A} \cdot \hat{x})(\hat{x} \cdot \hat{B}) + (\hat{A} \cdot \hat{y})(\hat{y} \cdot \hat{B}) \tag{30}$$

We already know how to write an expression for state vectors very much like Eq. 30. Figure 4 shows an experiment in which an input beam in state ψ is projected into state ϕ. The insertion of an analyzer loop does not change the experimental outcome but allows the expansion of $\langle\phi|\psi\rangle$ in the xy representation.

$$\langle\phi|\psi\rangle = \langle\phi|x\rangle\langle x|\psi\rangle + \langle\phi|y\rangle\langle y|\psi\rangle \tag{31}$$

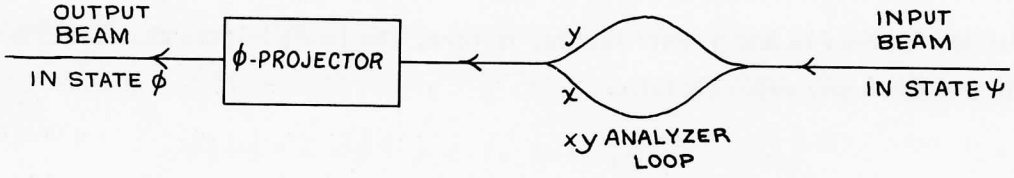


Figure 4. Experiment used to derive Equation 31.

Equation 31 has a form very similar to that of Eq. 30. Indeed, we can interpret Eq. 31 as an inner product between the dual vector $\langle\phi|$ and the vector $|\psi\rangle$. Suppose both $\langle\phi|$ and $|\psi\rangle$ are expanded in the xy basis (see Eqs. 3 and 19)

$$\begin{aligned} \langle\phi| &= \langle\phi|x\rangle\langle x| + \langle\phi|y\rangle\langle y| \\ |\psi\rangle &= |x\rangle\langle x|\psi\rangle + |y\rangle\langle y|\psi\rangle \end{aligned}$$

If we define inner product by the projection amplitude, the inner product of $\langle\phi|$ and $|\psi\rangle$ is

$$\begin{aligned} \langle\phi|\psi\rangle &= \Big[\langle\phi|x\rangle\langle x| + \langle\phi|y\rangle\langle y|\Big]\Big[|x\rangle\langle x|\psi\rangle + |y\rangle\langle y|\psi\rangle\Big] \\ &= \langle\phi|x\rangle\langle x|x\rangle\langle x|\psi\rangle + \langle\phi|x\rangle\langle x|y\rangle\langle y|\psi\rangle \\ &\quad + \langle\phi|y\rangle\langle y|x\rangle\langle x|\psi\rangle + \langle\phi|y\rangle\langle y|y\rangle\langle y|\psi\rangle \end{aligned} \tag{32}$$

Because of orthogonality between states x and y, this reduces to Eq. 31.

This second development of Eq. 31 justifies our speaking of the projection amplitude $\langle\phi|\psi\rangle$ as the inner product of a dual state vector $\langle\phi|$ and a state vector $|\psi\rangle$. The inner product in a representation other than xy can be derived by considering alternative experiments in which the xy analyzer loop in Figure 4 is replaced with an x'y' analyzer loop or with an RL analyzer loop.

We can summarize all such experiments and inner products between $\langle\phi|$ and $|\psi\rangle$ in all representations using the general expansions (16) and (26)

$$\langle\phi|\psi\rangle = \left(\sum_{j_i}\langle\phi|j_1\rangle\langle j_1|\right)\left(\sum_{j_2}|j_2\rangle\langle j_2|\psi\rangle\right) = \sum_{j_1}\sum_{j_2}\langle\phi|j_1\rangle\langle j_1|j_2\rangle\langle j_2|\psi\rangle \tag{33}$$

where the indices j_1 and j_2 run over the basis states of any representation (but the same representation for both sums).* The orthogonality of basis states of a given representation can be expressed by the relation

$$\langle j_1|j_2\rangle = \delta_{j_1 j_2} \begin{cases} = 1 \text{ if } j_1 = j_2 \\ = 0 \text{ if } j_1 \neq j_2 \end{cases} \tag{34}$$

Hence the double sum (33) reduces to a single sum

$$\langle\phi|\psi\rangle = \sum_j \langle\phi|j\rangle\langle j|\psi\rangle \tag{35}$$

In equation 35 the summation extends over any pair of basis states whatever. This demonstrates explicitly the invariance of the inner product under a change of representation.

The inner product can also be expressed using the row- and column-matrix notation.

$$\langle\phi|\psi\rangle = \begin{pmatrix}\langle\phi|x\rangle & \langle\phi|y\rangle\end{pmatrix}\begin{pmatrix}\langle x|\psi\rangle \\ \langle y|\psi\rangle\end{pmatrix} \tag{36}$$

Notice that the product of a one-by-two row matrix and a two-by-one column matrix <u>in that order</u> is a one-by-one matrix, i.e., a number. The invariance of the inner product is assured by the transformation laws (14), (22), etc., and the fact that the transformation matrices are unitary. For example, if we transform the row and column matrices in (36) to the RL representation according to Eqs. 15 and 23, we obtain

$$\begin{pmatrix}\langle\phi|R\rangle & \langle\phi|L\rangle\end{pmatrix}\begin{pmatrix}\langle R|\psi\rangle \\ \langle L|\psi\rangle\end{pmatrix} = \begin{pmatrix}\langle\phi|x\rangle & \langle\phi|y\rangle\end{pmatrix} U^{\dagger}(xy\rightarrow RL)\, U(xy\rightarrow RL)\begin{pmatrix}\langle x|\psi\rangle \\ \langle y|\psi\rangle\end{pmatrix} = \begin{pmatrix}\langle\phi|x\rangle & \langle\phi|y\rangle\end{pmatrix}\begin{pmatrix}\langle x|\psi\rangle \\ \langle y|\psi\rangle\end{pmatrix} \tag{37}$$

as may be verified by explicitly multiplying the U matrices.

Notice that the inner product involves one bra vector and one ket vector in the correct order to make a completed bracket. The order of the symbols in the bracket is significant, since $\langle\phi|\psi\rangle$ is not, in general, equal to $\langle\psi|\phi\rangle$ but is instead its complex conjugate. In a general vector space an inner product is defined between a vector and a dual vector. In the particular case of ordinary vectors the components of the dual of a vector A are the same as the components of A, so the order in the inner product does not matter.

* In general, in a multiple sum, we shall use j_1, j_2, etc., when the indices are to run over the basis states of the <u>same</u> representation. We will use i, j, k, etc., when the indices are to run over basis states of <u>different</u> representations.

5. Operators

We have seen in the preceding sections how a state vector provides an abstract description of a photon state, whereas row and column matrices provide a concrete description of that state in a particular representation. In Chapter 5 we showed that a two-by-two matrix provides a description of a device in a particular representation. We now introduce the concept of an operator, which describes a device independently of any representation.

Suppose a beam of photons in state ψ is incident on device A, and the output beam is monitored with a projector for some other state ϕ. Let open xy analyzer loops be inserted before and after device A as shown in Figure 5. These loops will not affect the outcome of the experiment. The probability amplitude that describes the experiment can be written in terms of the matrix elements of device A as follows:

$$\begin{aligned}\langle\phi|A|\psi\rangle = {} & \langle\phi|x\rangle\langle x|A|x\rangle\langle x|\psi\rangle + \langle\phi|x\rangle\langle x|A|y\rangle\langle y|\psi\rangle \\ & + \langle\phi|y\rangle\langle y|A|x\rangle\langle x|\psi\rangle + \langle\phi|y\rangle\langle y|A|y\rangle\langle y|\psi\rangle\end{aligned} \tag{38}$$

Since the arbitrary states ϕ and ψ are not necessarily basis states of the same representation, the amplitude $\langle\phi|A|\psi\rangle$ is not necessarily a matrix element of A in any single representation.

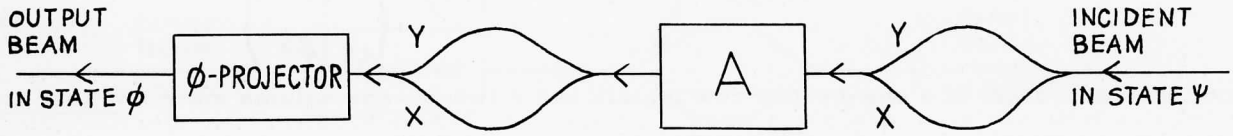


Figure 5. Experiment used to introduce the operator A.

We have seen in Section 4 that every projection amplitude can be thought of as the inner product of a dual vector and a state vector. Each term on the right side of Eq. 38 contains as a factor the inner product of $\langle\phi|$ with one of the basis vectors $|x\rangle$ or $|y\rangle$. The entire right side is therefore the inner product of $\langle\phi|$ and a linear combination of $|x\rangle$ and $|y\rangle$. The left side of (38) must also be the inner product of $\langle\phi|$ with something. We interpret the amplitude $\langle\phi|A|\psi\rangle$ as the inner product of $\langle\phi|$ with $A|\psi\rangle$, which represents a linear combination of state vectors. Thus we write

$$\begin{aligned}A|\psi\rangle = {} & |x\rangle\left[\langle x|A|x\rangle\langle x|\psi\rangle + \langle x|A|y\rangle\langle y|\psi\rangle\right] \\ & + |y\rangle\left[\langle y|A|x\rangle\langle x|\psi\rangle + \langle y|A|y\rangle\langle y|\psi\rangle\right]\end{aligned} \tag{39}$$

Here the expressions enclosed in square brackets stand for complex numbers.

Equation 39 describes the beam that emerges from device A when the input beam consists of photons in state ψ. But this state is arbitrary. We may formally delete the symbol $|\psi\rangle$ from the right hand end of every term in 39, and obtain the expression

$$\begin{aligned}A = {} & |x\rangle\langle x|A|x\rangle\langle x| + |x\rangle\langle x|A|y\rangle\langle y| \\ & + |y\rangle\langle y|A|x\rangle\langle x| + |y\rangle\langle y|A|y\rangle\langle y|\end{aligned} \tag{40}$$

Equation 40 may be regarded as merely a shorthand for Eq. 39. However, just as in the case of Eq. (3), one can attach a deeper significance to it. The symbol A on the left side of (40) can be considered an abstract operator that represents the device. The right side is a specific representation of this operator. An alternative representation of the same operator is obtained by replacing the xy analyzer loops in Figure 5 with RL analyzer loops. One then obtains by an identical procedure:

$$A = |R\rangle\langle R|A|R\rangle\langle R| + |R\rangle\langle R|A|L\rangle\langle L| + |L\rangle\langle L|A|R\rangle\langle R| + |L\rangle\langle L|A|L\rangle\langle L| \tag{41}$$

Equations (40) and (41) can be written compactly as double summations:

$$A = \sum_{i_1, i_2} |i_1\rangle\langle i_1|A|i_2\rangle\langle i_2| \tag{42}$$

If the dummy indices i_1, i_2 in the sums of Eq. 42 run over the values x, y, one gets Eq. 40 when the dummy indices run over the values R, L, one gets Eq. 41. But clearly any open analyzer loops could have been used in the experiment of Figure 5. Hence the indices in (42) can run over any complete set of basis states. Equation 42, with the basis states unspecified, provides a completely general representation of the operator A. When specific calculation is required, we must of course specify some particular representation.

What sort of mathematical object is an operator? Clearly it is not a number. Neither is it a state vector. It is still another entity, characterized by the following property: it acts on an arbitrary state vector to produce a linear combination of state vectors. An operator can also act on a dual vector to product a linear combination of dual vectors. Analogously to Eq. 39, one has the equation

$$\langle\phi|A = \langle\phi|x\rangle\langle x|A|x\rangle\langle x| + \langle\phi|y\rangle\langle y|A|x\rangle\langle x| + \langle\phi|x\rangle\langle x|A|y\rangle\langle y| + \langle\phi|y\rangle\langle y|A|y\rangle\langle y| \tag{43}$$

Notice that in the expansion of operator A (Eqs. 40 to 42), there appear matrix elements--which are numbers--and dual and state vectors written in that order. We have seen earlier that when written in the opposite order to produce a closed bracket a dual and a state vector define an inner product which is a number. The ket-bra $|\rangle\langle|$ placed back-to-back, where the blanks are filled by any two symbols for quantum states, represents an operator. The operator $|\alpha\rangle\langle\phi|$, for example, acting on the state vector $|\psi\rangle$ gives the state vector $|\alpha\rangle$ multiplied by the number $\langle\phi|\psi\rangle$. Alternatively, acting on the dual vector $\langle\beta|$, the same operator gives the dual vector $\langle\phi|$ multiplied by the number $\langle\beta|\alpha\rangle$. The product $|\rangle\langle|$ may be called an outer product of dual and state vectors. The outer product of dual and state vectors yields not a number nor another dual or state vector but an operator. In the notation of row and column matrices, observe that the product of a two-by-one column matrix and a one-by-two row matrix is a two-by-two matrix. This is just the matrix of the operator in the representation in which the row and column matrices are written.

Some examples will illustrate the procedure of representing operators. We consider some of the devices whose matrices were derived in Chapter 5.

Example 1. The identity operator

The matrix of an open hole or analyzer loop in any representation is the identity matrix (Equation 24 of Chapter 5)

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Use the xy representation and Eq. 40 to find the operator for the open hole or xy analyzer loop

$$I = |x\rangle\langle x| + |y\rangle\langle y| \quad \text{(open hole)} \tag{44}$$

Equation 44 is the machinery of an xy representation stripped to the running gears. Operating "from the left" on any state vector $|\psi\rangle$ it presents the state ψ in terms of the state vectors $|x\rangle$ and $|y\rangle$

$$I|\psi\rangle = |x\rangle\langle x|\psi\rangle + |y\rangle\langle y|\psi\rangle \tag{45}$$

Similarly, operating "from the right" on any dual vector $\langle\phi|$ it presents the state ϕ in terms of the vectors $\langle x|$ and $\langle y|$

$$\langle\phi|I = \langle\phi|x\rangle\langle x| + \langle\phi|y\rangle\langle y| \tag{46}$$

In terms of an arbitrary representation, the operator for the open hole or analyzer loop reduces Eq. 42 to a single summation:

$$I = \sum_j |j\rangle\langle j| \tag{47}$$

Example 2. Projection operators

From Eq. 40 and the known matrix elements of an x-projector (Eq. 16 of Chapter 5), we obtain the following expression for the operator that represents an x-projector:

$$P_x = |x\rangle\langle x| \tag{48}$$

We can verify directly from the definitions that (48) has the required properties. For if we operate from the left with (48) on $|x\rangle$ or $|y\rangle$, we get

$$P_x|x\rangle = |x\rangle\langle x|x\rangle = |x\rangle \tag{49a}$$

$$P_x|y\rangle = |x\rangle\langle x|y\rangle = 0 \tag{49b}$$

Operating on an arbitrary state vector $|\psi\rangle$, P_x gives

$$P_x|\psi\rangle = |x\rangle\langle x|\psi\rangle \tag{50}$$

Equations 49 and 50 epitomize the properties of an x-projector. Simiarly, we can verify that

$$P_R = |R\rangle\langle R| \tag{51}$$

represents an R-projector, andmore generally that

$$P = |\psi\rangle\langle\psi| \tag{52}$$

is a "projection operator" for the arbitrary state ψ. Operating on $|\psi\rangle$, P_ψ leaves it un-

changed; operating on a state vector for the state orthogonal to ψ, P_ψ gives zero; and operating on an arbitrary state vector $|\phi\rangle$ it gives $|\psi\rangle$ multiplied by the amplitude $\langle\psi|\phi\rangle$ whose absolute square measures the fraction of a beam in the state $|\phi\rangle$ that is transmitted by the ψ- projector.

Although a projection operator can always be written in the form (33), it may appear in a form which conceals this aspect of its nature. For example, an alternative expression for the operator P_R can be found using Eq. 3 of Chapter 5:

$$P_R = |x\rangle\,(1/2)\,\langle x| + |x\rangle\,(-i/\sqrt{2})\,\langle y| + |y\rangle\,(i/\sqrt{2})\,\langle x| + |y\rangle\,(1/2)\,\langle y| \qquad (53)$$

The numerical coefficients in (53) have been written between the state vector and the dual vector symbols in order to emphasize that this is a case of Eq. 40. Since a coefficient is a number, it can be put anywhere in its term.

Notice that according to Eq. 47, the identity operator that represents an open hole is just the sum of the projection operators for any pair of orthogonal basis states. Such a sum of projection operators gives the identity operator only if there is no relative phase factor between the two terms. For example, the operator

$$A = |x\rangle\langle x| - |y\rangle\langle y| \qquad (54)$$

is not the identity operator. It represents the device whose matrix in the xy representation is

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

The difference between (54) and (44) becomes even more marked when the operators are expressed in another representation. For example, let $|y'\rangle$ denote a photon state polarized at 45° to the state $|y\rangle$. Then the operator (54) can be written in the x'y' representation as

$$A = |x'\rangle\langle y'| - |y'\rangle\langle x'| \qquad (55)$$

This is not a sum of projection operators, whereas I is a sum of projection operators in all representations. Equation 55 makes it manifest that the device is one that turns x' polarized beams into y' polarized beams, and vice versa.

6. Summary

In all the examples above, device A acts on an input beam to produce an output beam. Corresponding to this action of the device on the beam is the action of the operator on the state vector that describes the input beam. In general, the effect of operator A on an arbitrary state vector is to change it to some other state vector (or linear combination of state vectors) and multiply it by a numerical constant. The absolute square of this constant specifies the intensity of the output beam relative to that of the input. Operators, and dual and state vectors, taken together, provide a summary notation for describing the effects of devices on beams.

APPENDIX TO CHAPTER 6 *

TWO-PHOTON STATES AND THE ANNIHILATION OF POSITRONS

1. Introduction

Our discussion thus far has dealt entirely with one-photon states. It was emphasized in Chapter 3 that even though all the measurements described involve beams of photons, the photons interact with the measurement apparatus one at a time. Different photons in the beam act entirely independently of one another.

Sometimes two photons are emitted from a single atomic or nuclear event, such as the decay of a particle or the mutual annihilation of a particle-antiparticle pair. In such a case the two photons involved are correlated with one another and they must be described by a two-photon state. In the present appendix we extend the idea of quantum state to include states of two photons. Such concepts as orthogonality, completeness, change of basis, and the like apply to these states. Determination of a two-photon state requires the study of many pairs of photons, just as the determination of a one-photon state requires the study of many single photons. The analysis of two-photon states presented here can readily be extended to describe states of three or more photons.

2. Positron-electron annihilation

An introduction to two-photon states is provided by the study of positron-electron annihilation. Positrons are given off in the radioactive decay of certain nuclei (for example, copper 64 and sodium 22). When the radioactive nuclei are embedded in a solid, many of the positrons are brought to rest in the solid and thereafter annihilate with electrons of the solid to produce gamma rays. The characteristic time for a positron-electron annihilation is about 10^{-10} seconds. In a solid the positron is brought to rest, on the average, in a time approximately one thousand times shorter than this. Therefore most annihilations take place from rest.

We are interested here in a particular type of annihilation event, namely one in which two gamma rays are produced

$$e^+ + e^- \longrightarrow \gamma + \gamma \qquad (1)$$

Reaction (1) is the most frequent mode of annihilation. (Notice that the emission of a single gamma ray from an isolated electron-proton system is forbidden by energy-momentum conservation.) The two-photon decay can be recognized by the fact that the photons always come off in

---

* This is a preliminary version of what will eventually be a separate chapter. The authors acknowledge the advice of Professor Stephan Berko in preparing this appendix.

opposite directions, each with an energy 0.511 MeV, the rest energy of an electron or positron.

Because the electron and positron annihilate from rest, they are most likely to have zero angular momentum. This intuitive classical picture is corroborated by a detailed quantum mechanical analysis. As a result, the two emitted gamma rays must carry zero total angular momentum. This condition restricts the possible polarization states of these photons, as can be seen most directly by analyzing the states in terms of circular polarization. Suppose an L-polarized photon is emitted in some direction from an annihilation. As we saw in Chapter 2, this photon carries angular momentum $+\hbar$ along its direction of motion. Both linear and angular momentum are conserved if another L-polarized photon is emitted in the opposite direction. (Since the directions of motion of the gamma rays are opposite, the angular momenta will then also be opposite and will add up to zero.) Likewise both linear and angular momentum are conserved if two R-polarized photons (each with angular momentum $-\hbar$ along its direction of motion) are emitted in opposite directions. Angular momentum would not be conserved if one of the emitted photons were right-circularly polarized and the other left-circularly polarized. In summary, we expect that when the emitted pairs of gamma rays are tested for circular polarization, the polarization will be sometimes RR and other times LL but never RL or LR.

An experiment that would confirm the conclusions of the preceding paragraph is sketched in Figure 1. Circular polarization analyzers are placed on opposite sides of the source as

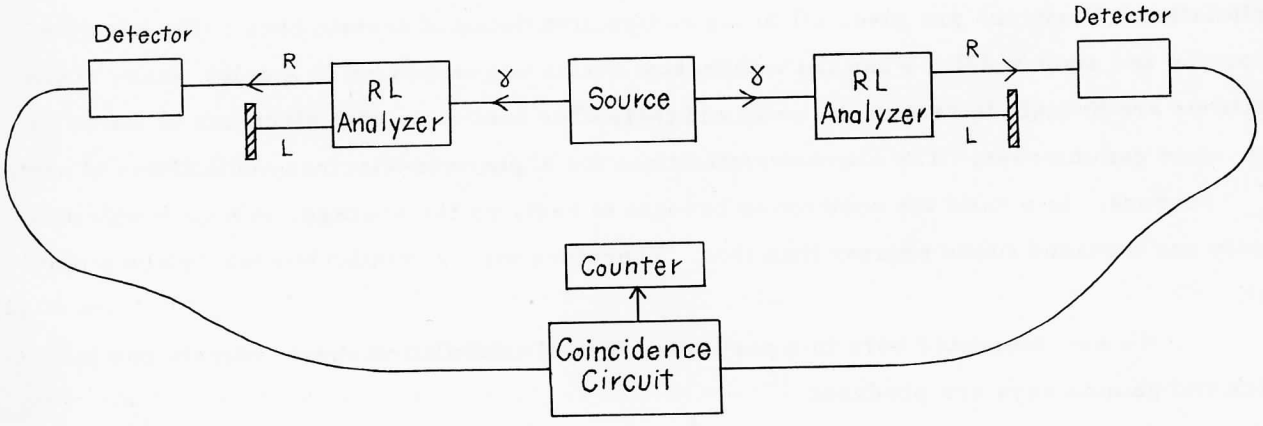


Figure 1. Idealized experiment to verify that γ-ray pairs from decay of positronium are polarized RR or LL but not LR or RL.

shown. The outputs of the detectors go to a coincidence circuit, which records a count only when it receives simultaneously (within a very short time interval) a signal from each input.

The argument based on the assumption that the initial state has zero angular momentum implies that the coincidence circuit should register counts when both detectors are connected to the R output channels of the analyzers or when both are connected to the L channels. But no coincidence counts should be recorded when one side is connected to an R channel and the other to an L channel.

It happens that the experiment of Figure 1 has not been carried out for positron-electron annihilation. It is not easy to test gamma rays for circular polarization. At the photon energies involved neither calcite nor quartz is sufficiently birefringent and quarter-wave plates cannot be constructed. It is possible to measure circular polarization of gamma rays using magnetized iron, although the efficiency of detection is low. Circular polarization has been detected with this technique in other experiments.* In the case of positron-electron annihilation, no one has yet taken the trouble to perform the rather difficult circular polarization experiment.

Linear polarization of gamma rays is much easier to measure than is circular polarization. The technique is based on the different scattering coefficients for gamma rays of different linear polarization when they are incident on a target. A similar experiment to determine the linear polarization of visible light could be carried out by measuring its reflectivity from a dielectric surface (Section 2 of Chapter 3). In Section 4 of this chapter we describe an experiment for measuring the linear polarization of gamma rays from positron-electron annihilation and show how the results of this experiment can be used to obtain information concerning the state of the two-photon system. First we discuss in Section 3 the formal description of two-photon states.

3. Two-photon states

The notation for describing two-photon states is a straightforward extension of that introduced earlier for single-photon states. For example, if we have photon pairs which are all right-circularly polarized, we symbolize this two-photon state by the state vector $|RR\rangle$. In like manner we define the state vectors $|LL\rangle$, $|LR\rangle$, and $|RL\rangle$. The dual vectors of all these states are similarly defined. We can deduce from our knowledge of single-photon states that the four states $|RR\rangle$, $|LL\rangle$, $|RL\rangle$, and $|LR\rangle$ constitute a complete set for the two-photon system. To show this let $|\psi\rangle$ denote the two-photon state of a pair of photons emitted in positron-electron annihilation or in any other process. Then the coincidence rate in the idealized experiment of Figure 1, divided by the coincidence rate with the analyzers absent, measures the projection probability $|\langle RR|\psi\rangle|^2$. Similar experiments measure $|\langle RL|\psi\rangle|^2$, $|\langle LR|\psi\rangle|^2$ and $|\langle LL|\psi\rangle|^2$. The sum of these four quantities must be unity for the simple reason that each photon, taken alone, must emerge in one or the other channel of an RL

* See article by M. Goldhaber, L. Grodzins, and A. W. Sunyar in Alpha, Beta, and Gamma Ray Spectroscopy Edited by K. Siegbahn, North Holland Publishing Co., 1965, pp. 1423-1431.

analyzer. Therefore the set of states is complete. We can treat each photon separately in this manner because the two photons in a two-photon state do not interact once they have been emitted in opposite directions. Measurement of one of these photons does not affect the results of experiments with the other member of the pair.

Not only do the four two-photon basis states form a complete set; they are also mutually orthogonal. For example, in the experiment of Figure 2, which measures the projection probability $|\langle RL|\,RR\rangle|^2$, the coincidence counting rate must be zero.

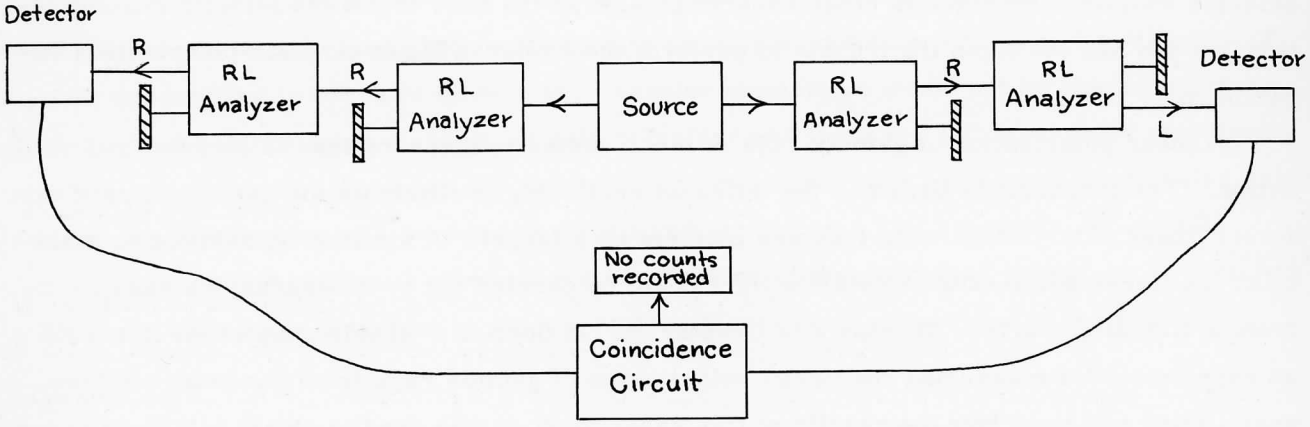


Figure 2. Idealized experiment to show that state $|RL\rangle$ is orthogonal to state $|RR\rangle$ as predicted from results with photons in the visible region.

In an entirely analogous manner we define a set of four linear polarization basis states $|xx\rangle$, $|xy\rangle$, $|yx\rangle$, and $|yy\rangle$. These states are also mutually orthogonal and form a complete set. The absolute squares of the amplitudes $\langle xx|\psi\rangle$, $\langle xy|\psi\rangle$, $\langle yx|\psi\rangle$, and $\langle yy|\psi\rangle$ are measured by experiments similar to that of Figure 1, with linear polarization analyzers replacing the circular polarization analyzers. These amplitudes must be related to $\langle RR|\psi\rangle$, $\langle RL|\psi\rangle$, etc. by a transformation matrix whose elements are the projection amplitudes $\langle xx|RR\rangle$, $\langle xy|RL\rangle$ and so on. The latter amplitudes represent the results of experiments in which pairs of circularly polarized photons are passed through appropriate linear polarization projectors. Since the two photons of a pair do not interact, measurements on one photon in the pair does not affect the other photon. Therefore the two-photon projection amplitudes must be products of one-photon amplitudes,

$$\begin{aligned} \langle xx|\,RR\rangle &= \langle x|\,R\rangle\langle x|\,R\rangle \\ \langle xy|\,RL\rangle &= \langle x|\,R\rangle\langle y|\,L\rangle \end{aligned} \qquad (2)$$

and so on. A wide variety of experiments lead us to believe that photons of all frequencies have fundamentally the same polarization properties. Therefore we can construct the full transformation matrix for gamma ray two-photon states from our knowledge of probability amplitudes for photons in the visible region (Table 1 of Chapter 4).

$$U = \begin{pmatrix} \langle xx|RR\rangle & \langle xx|RL\rangle & \langle xx|LR\rangle & \langle xx|LL\rangle \\ \langle xy|RR\rangle & \langle xy|RL\rangle & \langle xy|LR\rangle & \langle xy|LL\rangle \\ \langle yx|RR\rangle & \langle yx|RL\rangle & \langle yx|LR\rangle & \langle yx|LL\rangle \\ \langle yy|RR\rangle & \langle yy|RL\rangle & \langle yy|LR\rangle & \langle yy|LL\rangle \end{pmatrix}$$

$$= \begin{pmatrix} -1/2 & 1/2 & 1/2 & -1/2 \\ -i/2 & -i/2 & i/2 & i/2 \\ -i/2 & i/2 & -i/2 & i/2 \\ 1/2 & 1/2 & 1/2 & 1/2 \end{pmatrix} \quad (3)$$

Of course the phases in the elements of U are determined by the phase conventions chosen in Chapter 4.

A decomposition of the form (2) is not possible if the two particles interact at the time of measurement. Later on we shall encounter many examples of two-particle amplitudes which do not separate in this simple manner, for instance that of the electron and proton in the hydrogen atom. However, the expansion of an arbitrary state vector in terms of basis vectors, and the existence of a transformation matrix like (3) between one basis and another, are quite general features of the analysis of two-particle systems. Such procedures can be further generalized to apply to systems containing any number of particles. The number of states in the basis and the labels used to describe the states depend, of course, both on the number and the type of particles in the system. If we are concerned only with polarization states of photons, a system of n photons has a basis consisting of 2^n states.

4. The state of photons from positron annihilation.

We return now to the analysis of the state of the gamma ray pairs emitted in positron-electron annihilation. We have surmised, on the basis of angular momentum conservation, that the amplitudes $\langle RL|\psi\rangle$ and $\langle LR|\psi\rangle$ both vanish. Symmetry considerations suggest (and the experiment described below confirms) that the projection probabilities $|\langle RR|\psi\rangle|^2$ and $|\langle LL|\psi\rangle|^2$ are equal. Each probability then has the value 1/2 and the two-photon state vector may be written in the form

$$|\psi\rangle = |RR\rangle(e^{i\alpha}/\sqrt{2}) + |LL\rangle(e^{-i\alpha}/\sqrt{2}) \quad (4)$$

Here the absolute phase--which has no physical significance--has been chosen for convenience. The value of the phase α is determined from the results of a linear polarization experiment.

With the help of the transformation matrix (3), we can write the state vector (4) in terms of linear polarization basis states

$$|\psi\rangle = |xx\rangle(-\cos\alpha)/\sqrt{2} + |xy\rangle(\sin\alpha)/\sqrt{2} + |yx\rangle(\sin\alpha)/\sqrt{2} + |yy\rangle(\cos\alpha)/\sqrt{2} \quad (5)$$

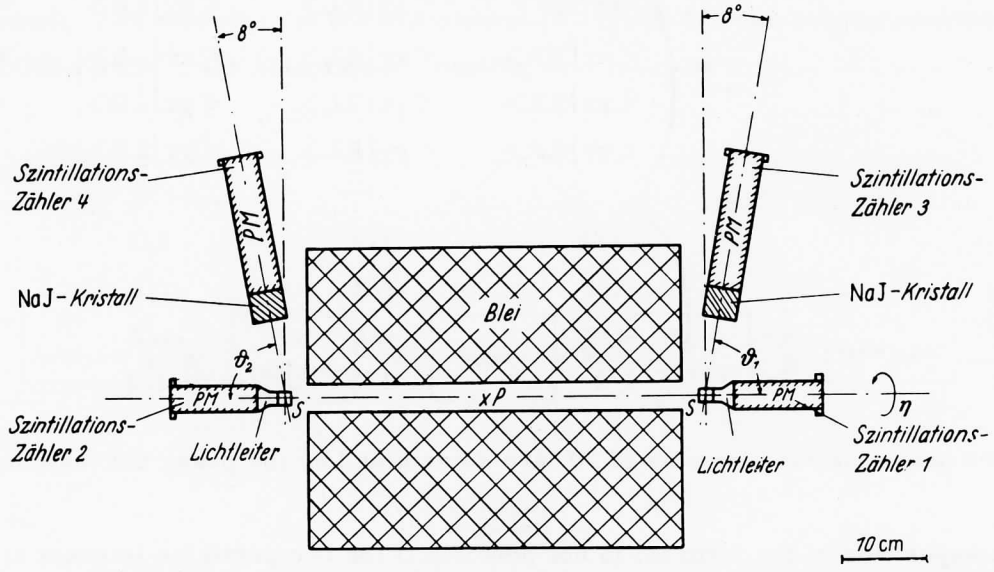


Fig. 1. Versuchsaufbau. Die vom Positronenpräparat P ausgehenden Vernichtungsquanten werden an den als Streukörper dienenden Szintillatoren S um den Winkel ϑ gestreut und dann in den NaJ-Kristallen absorbiert. Die vier mit PM bezeichneten Photomultiplier wandeln die Lichtblitze in elektrische Impulse um

Figure 3. Illustration from article by Langhof. German-English dictionary: Blei = lead, Szintillationszähler = scintillation counter.

Measurement of the absolute squares of the amplitudes $\langle xx|\psi\rangle$, $\langle xy|\psi\rangle$ etc. will fix the value of the unknown phase α. This measurement was first carried out by Wu and Shaknov (Physical Review, 77, 136 (1950)). A refined version of the experiment was later performed by Langhoff.* Figure 3 shows Langhoff's experimental arrangement. The source of positrons is located in the center of a block of lead. A straight hole drilled through the lead permits the exit of gamma rays moving in opposite directions from the source. The emerging gamma rays are detected using scintillation crystals. A scintillation crystal emits a flash of light when a gamma ray passes through it. This flash of light is detected by a photomultiplier tube fixed to the crystal. A few of the incoming photons are Compton scattered by electrons in the first crystal S and are detected by a second scintillation counter (either counter 3 or counter 4 in the figure). A single gamma ray detected by both counter 1 and counter 3, for example, must have been scattered by electrons in crystal S through an angle θ_2 equal to 82°.

A classical argument suggests why the arrangement of scintillation counters in Figure 3 acts as a projector for a particular linear polarization. In a classical description, the elec-

* H. Langhoff, Zeitschrift für Physik, 160, 186 (1960).

tric field of the incident wave sets an electron in the crystal into vibration. This vibration gives rise to radiation in all directions, which is observed as a scattered wave. The intensity of the radiation emitted in a particular direction is proportional to the square of the component of acceleration transverse to the direction of observation. Suppose one observes the radiation scattered at 90° to the incident direction. For detector A in Figure 4, the vibration of the electron is transverse to the direction of observation and a scattered wave is detected. However, for detector B in Figure 4, the electron vibration is entirely along the direction of observation and no scattered wave will be observed. Consequently, if the 90° scattering is observed the system ought to function as a projector for linear polarization. That is, the reflected beam is polarized along the axis perpendicular to the plane determined by the incoming and outgoing directions.

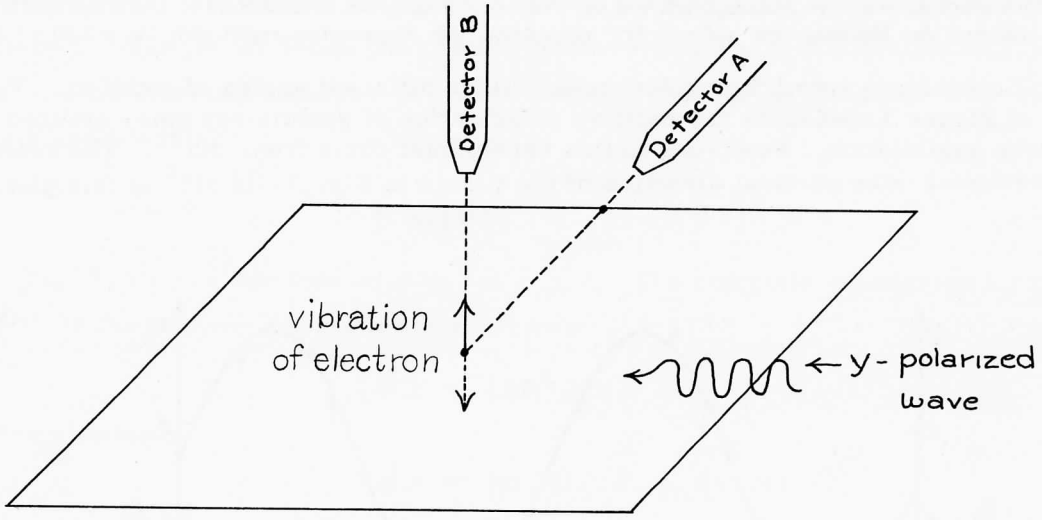


Figure 4. Classical analysis of detection of linear polarization of gamma rays by Compton scattering. The intensity at detector A is much greater than the intensity at detector B.

Looked at as a quantum phenomenon, the scattering of gamma rays by the electrons in a solid is primarily Compton scattering (Chapter 2, page 2-5). A quantum treatment of Compton scattering in a solid yields results for the polarization of the scattered gamma rays similar to that of the preceding oversimplified classical discussion. The principal differences are, first, that the scattering angle for maximum polarization is 82° rather than 90° due to the recoil of the scattering electron (for gamma ray energy 0.511 MeV--see the exercises); and, second, the polarization is not one hundred percent even at the angle of maximum polarization.

The purpose of the experiment of Figure 3 is to relate the polarizations of two oppositely-moving gamma rays that originate from a single positron-electron annihilation. For this purpose simultaneous scattering events are detected at both ends of the hole in the lead block. Counts are recorded only when all four scintillation counters emit signals at the same time (within some small prescribed time interval). Let the x-axis point out of the page and the y

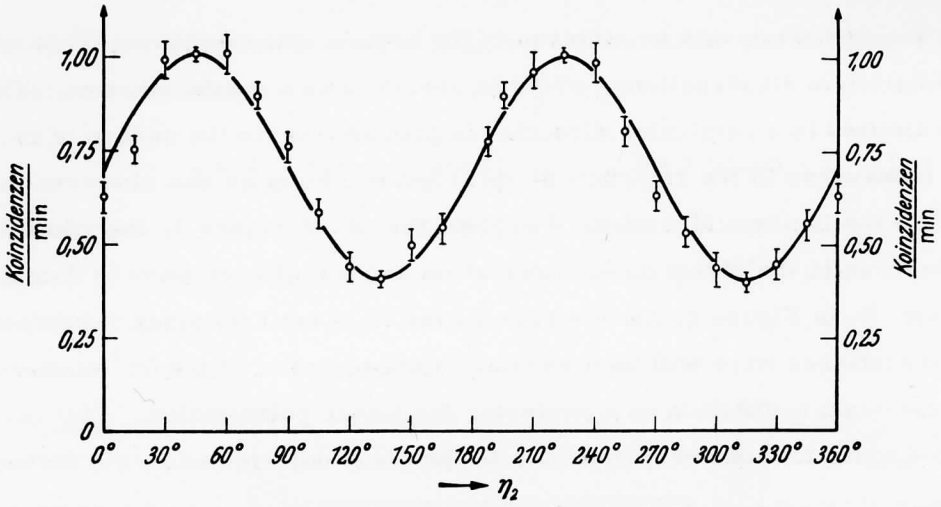


Fig. 3. Ergebnis der Polarisationsmessungen mit Na^{22}. Aufgetragen ist die Zahl der pro Minute stattgefundenen Vierfachkoinzidenzen in Abhängigkeit von der Winkelstellung η_2 des Polarimeters 2. Das Polarimeter 1 blieb während der Messung fest auf $\eta_1 = 315°$ eingestellt. Als Asymmetrie ergibt sich $U_{ex} = 2{,}50 \pm 0{,}10$

Figure 5. Coincidence rate ("Koinzidenzen/min") for different angles of rotation η in the apparatus of Figure 3 measures the relative polarization of gamma ray pairs emitted in positron-electron annihilation. Positrons in this experiment come from Na^{22}. The reference angle of scattering--the vertical direction of the y-axis in Fig. 3--is 315^{o} in this plot.

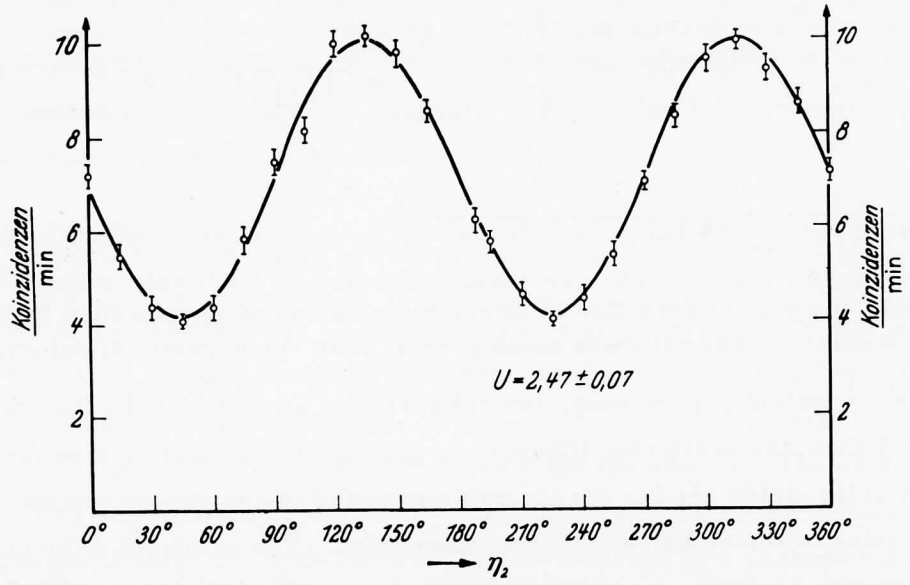


Fig. 4. Ergebnis der Polarisationsmessungen mit Cu^{64}. Aufgetragen ist die Zahl der pro Minute stattgefundenen Vierfachkoinzidenzen, bezogen auf 100 mC Cu^{64}, in Abhängigkeit von der Winkelstellung η_2 des Polarimeters 2. Das Polarimeter 1 war während der Messung in der 45°-Stellung. Der Untergrund an „falschen" Koinzidenzen von 4,2% wurde bereits abgezogen. Als Asymmetrie ergibt sich $U_{ex} = 2{,}47 \pm 0{,}07$

Figure 6. Same as Figure 5, using Cu^{64} as a source of positrons. Reference angle of scattering is 45^{o}. Similarity of the results shown in Figures 5 and 6 is evidence that gamma ray pairs are characteristic of positron-electron annihilation and not of source of positrons.

axis toward the top of the page. Then the experimental arrangement shown in Figure 3 constitutes a measurement of the probability $|\langle xx|\psi\rangle|^2$. Counter number 3 on the right hand end can be rotated about an axis given by the gamma ray beam (the angle shown as η in the figure). When this angle has been increased by 90 degrees, the experiment will measure the absolute square of the amplitude $\langle xy|\psi\rangle$. The results of the experiment are plotted in Figures 5 and 6 for two different positron sources; the positrons emitted by the two sources originate from different nuclear decays and have different initial energies. The similarity of the two figures is evidence that the gamma rays come from positron-electron annihilation and not from some other reaction involving the original positrons.

In both Figure 5 and 6 one sees a maximum in the coincidence counting rate when the analyzing crystals are oriented to detect orthogonal polarizations and a minimum when they are oriented to detect the same polarization. The minimum in the latter case is not zero because there is no angle for which the crystal scatters only a single polarization. Further theoretical and experimental analysis provides evidence that for an ideal polarizer the minima of Figures 5 and 6 would be zero. In other words

$$|\langle xx|\psi\rangle|^2 = |\langle yy|\psi\rangle|^2 = 0 \qquad (6)$$

Looking at Eq. 5, we see that this implies $\alpha = \pi/2$. The complete expression for the state vector $|\psi\rangle$ in the RL representation is therefore

$$|\psi\rangle = |RR\rangle(i/\sqrt{2}) - |LL\rangle(i/\sqrt{2}) \qquad (7)$$

and equation 4 reduces to

$$|\psi\rangle = |xy\rangle(1/\sqrt{2}) + |yx\rangle(1/\sqrt{2}) \qquad (8)$$

A similar coincidence experiment has been performed on the decay of the neutral pi meson (π^o), which likewise can decay into two gamma rays. These experiments give the same results (7) and (8) as for positron-electron annihilation. This experiment gives valuable evidence concerning the so-called "intrinsic parity" the neutral pi meson. Because of the minus sign in Eq. 7, the intrinsic parity is said to be negative.

5. Necessity for a two-photon description

The central feature of two-photon states is the correlation between pairs of photons. This correlation cannot be detected by making observations on only one photon of the pair. For example, taken by itself, there is no polarization of photons in the beam emerging from the right hand end of the lead block in the experiment of Figure 3. Suppose we record only coincidences between the two counters (counters 1 and 3) at this end. Then the coincidence rate is uniform for all linear polarizations--all angles of rotation η of counter 3. Moreover, we expect that in a modified experiment we could detect emerging from the right hand end both right- and left-circularly polarized photons in equal numbers. In brief, the photons emerging from the right

hand end of the hole are unpolarized according to the definition for one-photon states (Section 10 of Chapter 3). The same must be true of the photons emerging from the left end of the hole, when these photons are considered in isolation. There is no unique orientation of counter 4 with regard to rotation about the beam direction. All such orientations give uniform coincidence rates for the linear polarization of gamma rays emerging from the left end. Therefore we can with equal validity choose the y-axis--coincident with the plane in which counter 4 lies--to point in any direction transverse to the beam.

Polarization effects dependent on angle occur only when analyzers and detectors are used at both ends of the hole through the lead block. Then we observe relative polarizations of pairs of oppositely-moving gamma rays. It is this feature of correlation that makes necessary a two-photon description.

Chapter 7. Atomic Beams

1. Introduction

Our study of photon polarization states in Chapters 3 through 6 has revealed many of the central features of quantum physics. For photons these features can be exhibited in tabletop experiments employing relatively simple equipment. With the present chapter we begin a study of atomic and sub-atomic particles that have non-zero rest mass. The concepts and formalisms developed in the study of photon polarization provide the basis for understanding the behavior of these particles. The experiments that demonstrate the properties of these particles typically cannot be performed on a tabletop but most often require a good deal of complicated equipment. Moreover, there is no classical theory, analogous to wave optics, that prepares us for the results of the experiments.

We start with a class of experiments involving "atomic beams." This generic term covers beams of various types of particles including atoms, molecules, and neutrons, all of which share the common property that they are electrically neutral. The quantum states that describe such systems involve many variables. Just as the photon experiments that we studied provide information on only one part of the complete photon state (the polarization properties), the atomic beam experiments single out one aspect of the quantum nature of the systems under study--the so-called "spin states." Spin states involve the magnetic dipole moment and angular momentum of the atom or molecule. Other aspects of atomic quantum states will be studied in later chapters.

The experiments to be described lead to the conclusion that the component of the atomic angular momentum along any fixed axis can assume only certain discrete values. Each of these values defines a quantum state. A complete orthogonal set of basis states for the description of the angular momentum or spin consists of states with all the allowed components with respect to a specified fixed axis, called the quantization axis. For different atoms a complete set may contain two, three, four, or more states. The mathematical description of these states constitutes a simple extension of the formalism developed for the two-state photon system; a spin state is uniquely defined by its projection amplitudes or components with respect to some specific basis. Change of basis involves a change in the direction of the quantization axis, and is governed by transformation laws analogous to those for photon states. The nature of the transformation depends only on the number of states in a complete set.

2. Atomic beam apparatus

In a typical atomic beam experiment particles are produced in a source, pass through one or more regions in which they undergo an interaction with externally-applied fields, and finally enter a detector to be counted. The purpose of such an experiment is to study the properties of isolated atoms; it is clearly important for this purpose that the atoms not collide either with one another or with air molecules during their flight between source and detector. Most beams from actual sources are not so intense as to make collisions among beam particles an important problem. However, at atmospheric pressure and room temperature the mean free path for collisions with air molecules is only about 10^{-4} cm. Consequently the experiments must be carried out at pressures of the order of 10^{-6} atmospheres or lower. Figure 1 is a picture of an atomic beam apparatus. Clearly we have left behind the simple experiments with which we were able to study polarized photons: an atomic beam experiment requires a vacuum chamber with its associated pumps, valves, and gauges, appropriate sources and detectors, as well as magnets and other specialized equipment to provide interaction with the beam.

The central fact of electrical neutrality governs the design of sources and detectors for atomic beams. If atoms in the beam were charged, it would be relatively easy to collimate them by means of electric and magnetic fields, and to detect them by virtue of the charges they carry. But it would also be extremely difficult to separate such ionic beams according to their quantum states in the manner we are about to describe. Therefore one uses neutral beams, and neutral atoms lack the handle of electric charge by which they can be prepared, manipulated, and counted. Acceleration to a well defined high speed, focusing, and detection are all difficult to accomplish.

Some atomic beam sources are listed in Table 1. Neutrons are produced in a reactor. Gas molecules may be allowed to stream through an aperture. Metals can be vaporized in an oven and the atoms of vapor released via an aperture. If diatomic molecules are to be dissociated to permit study of the atoms individually, the energy of dissociation can be supplied by electric discharge or heating.

A spreading beam from the source can be masked down with slits to produce a narrower beam*. The velocities of the particles in the beam vary over a range that depends upon the temperature of the source. If a nearly mono-energetic beam is required, a series of shutters which open and close in sequence can be used to eliminate atoms moving at velocities other than the desired one. Unfortunately this procedure reduces considerably the intensity of the resulting beam.

The beam itself is not visible to the human eye, and must be "observed" by means of a suitable detector (Table 2) whose output is read on a meter or chart recorder. The kind of detector most used ionizes the atom and measures the ion current that results. Neutrons of course cannot be ionized, and other techniques must be employed to detect them.

* But the beam cannot be made arbitrarily narrow (See Chapter 9).

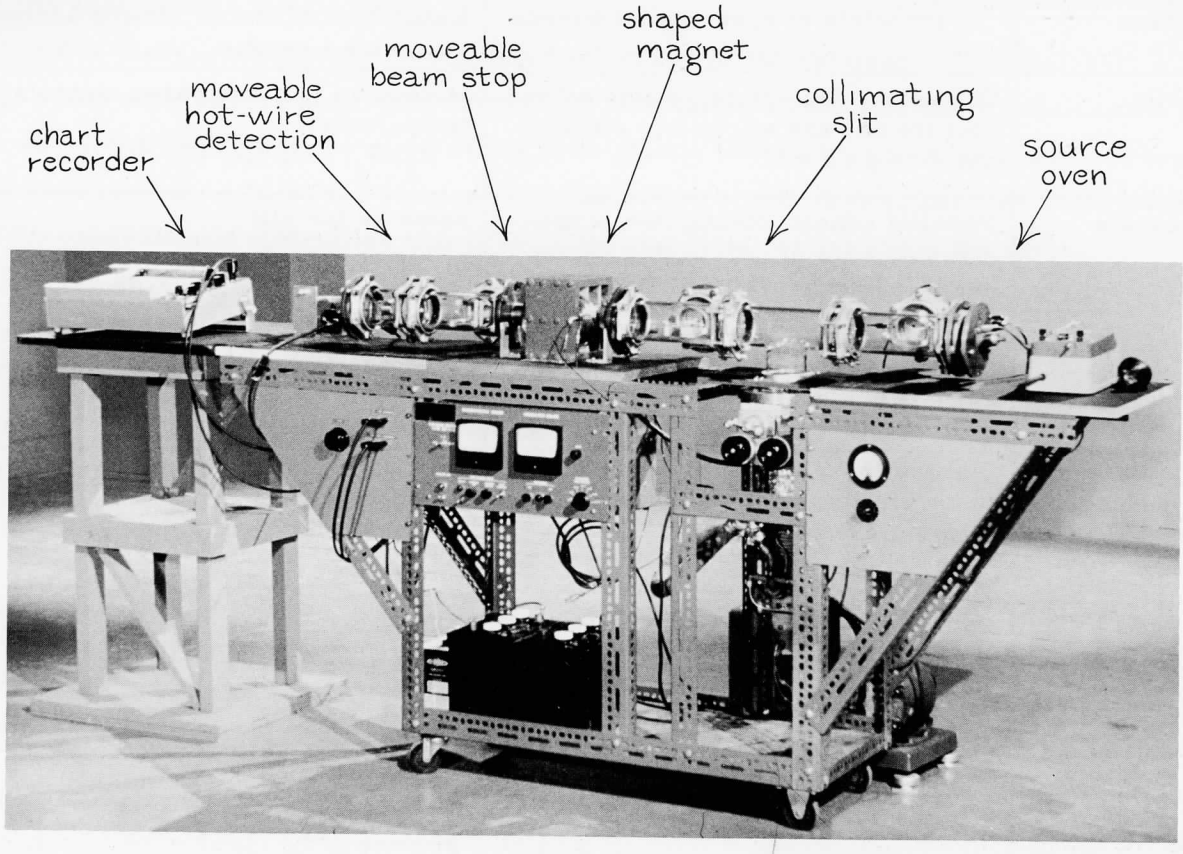


Figure 1. Atomic beam apparatus for displaying spin states of cesium. From the film "The Stern-Gerlach Experiment" produced by Educational Services, Inc.

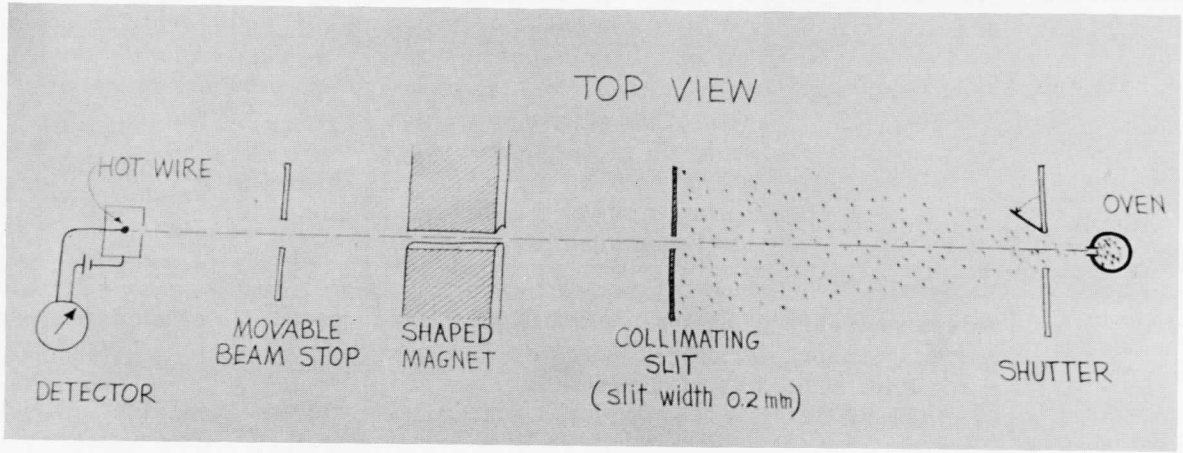


Figure 2. Schematic diagram of the atomic beam apparatus pictured in Fig. 1. Operation of the shaped magnet and hot wire detector are explained in the text. Width of oven slit: 0.12 mm.

Table 1. Some types of atomic beam sources*

Source	Principle of operation of source	Examples of "atoms" emitted from such a source
slit	Gas molecules are piped into an enclosure and are allowed to stream out through a slit	Molecules that are gases at room temperature: N_2, H_2, He, Ne, O_2, Cl_2 ....
canals	Parallel canals looking like honeycomb or edge of corrugated cardboard. Canals afford greater directionality than slits. To increase beam flux use many parallel canals.	same as for slit
oven (used with slit or canals)	vaporizes non-gaseous elements	Li, B, Na, K, Cu, Cs, Au ....
electrical discharge tube, microwave arc, dissociating oven (used with slit)	dissociates polyatomic gases into monatomic gases by electrical or thermal excitation.	monatomic gases: H, N, Cl, O....
Reactor	Controlled chain reaction	neutrons

Table 2. Some types of atomic beam detectors*

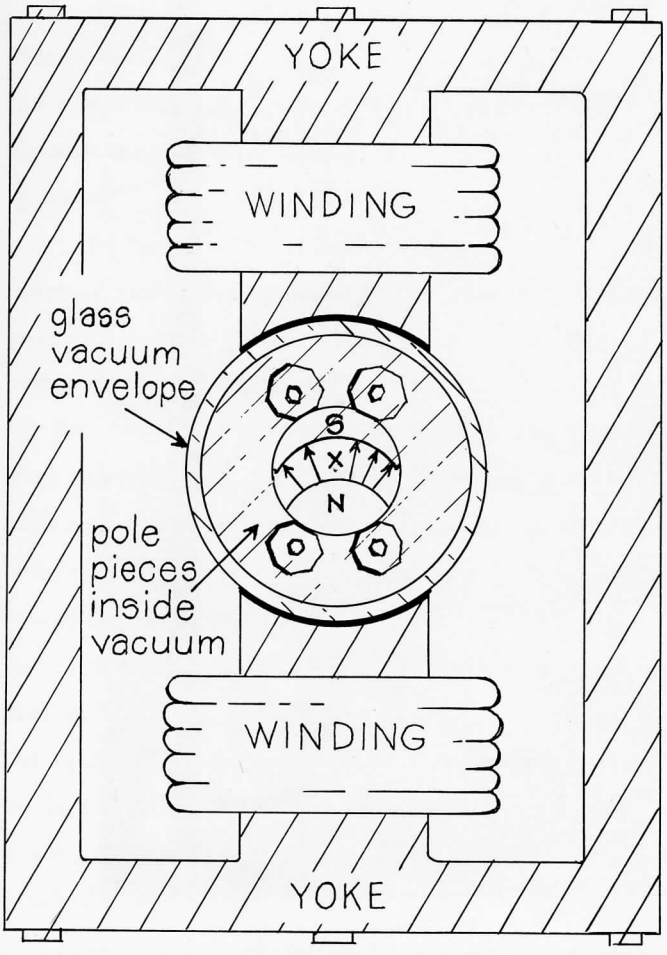
Type of detector	Principle of operation of detector	Examples of "atoms" detected by this method
deposition detector	Beam falls on a target surface which holds atoms or reacts chemically with them. Deposit may be detected by sight, with or without chemical development, or (if atoms are radioactive) by nuclear counting	Ag, Bi, Hg, K .... H, O, N .... especially useful for radioactive atoms
ionizing detectors	Converts neutral atoms in beam to ions by several alternative means: (a) When neutral atoms fall on a hot wire some lose an electron making positive ion (b) other atoms will receive an additional electron, making negative ion. (c) A stream of electrons crossing the beam will ionize some of the atoms. Resulting ions may be analyzed with a mass spectrometer; resulting electrons may be detected with an electron multiplier.	(a) easily-ionized atoms: Na, K, Cs .... (b) F, Cl, Br .... (c) any atoms whatever ("universal ionizer").
Pirani gauge	Beam enters a cavity by a single channel, raises pressure in cavity. Hot filament in cavity is cooled more by higher pressure gas. Rate of cooling of filament is measured electrically.	now obsolescent; very important in older experiments reported in the literature
BF_3 counter	neutron capture in Boron 10 makes an energetic nuclear reaction that produces detectable ionization	Neutrons

* See "Some New Applications and Techniques of Molecular Beams" by John G. King and Jerrold R. Zacharias in Volume VIII of Advances in Electronics and Electron Physics, L. Marton, ed. Academic Press, Inc., New York, New York, 1956.

3. An analyzer for atomic beams

Our discussion of photon states has been based on the existence of an analyzer, a device that splits an arbitrary beam into components that define a complete set of polarization states. An analogous device defines those states of an atomic beam with which we are now concerned. We first present an idealized version of the operation of this device; later we shall discuss some of the experimental difficulties that prevent the complete realization of the ideal.

The analyzer for atomic beams makes use of an inhomogeneous magnetic field to deflect the atoms from their straight-line path. Such an experimental arrangement was first employed by Stern and Gerlach in 1922. We shall define as a "standard analyzer" the magnet sketched in Figure 3; notice that the field is stronger near the north pole. When the device is oriented with its N-S axis in the z-direction, we call it a z-analyzer. The field B then points essentially



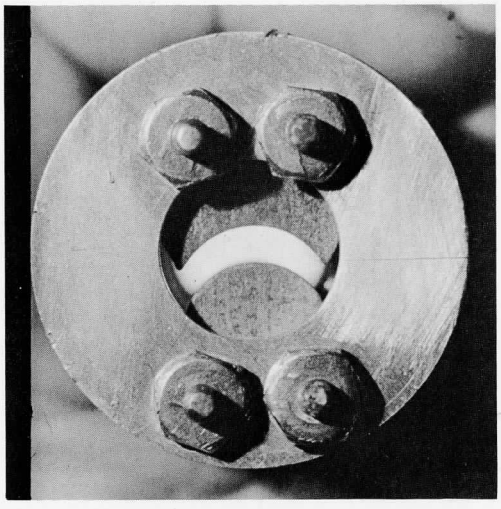


Figure 3. Cross-section of magnet used in atomic beam apparatus of Fig. 1. Beam travels through location marked x in a direction perpendicular to the page. Photo at right is of pole pieces from vacuum chamber.

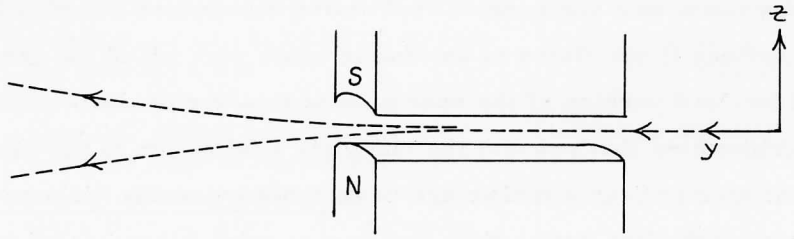


Figure 4. Idealized action of a magnetic analyzer on a two-state spin system. The magnitude of the beam deflections is much exaggerated.

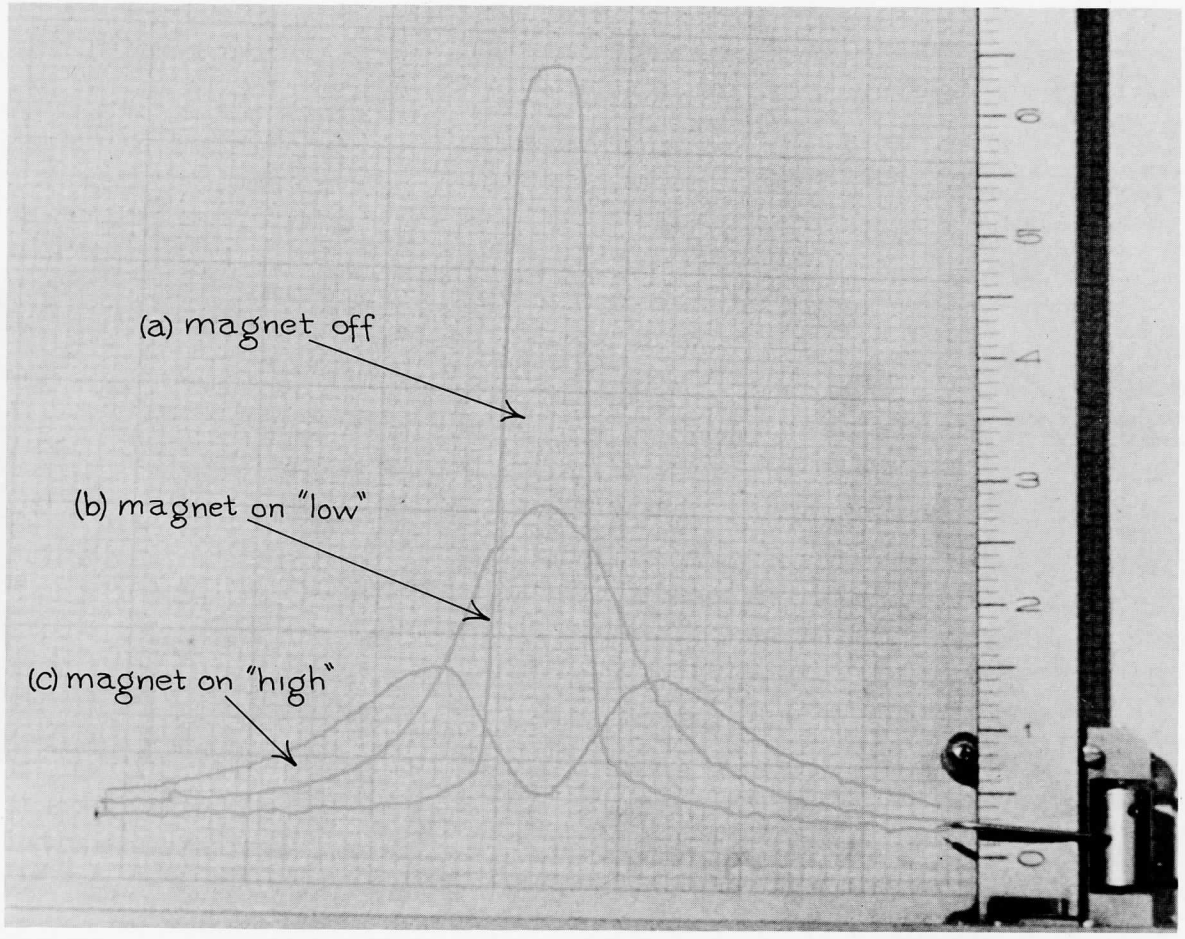


Figure 5. Beam profiles obtained for a cesium beam using the equipment of Figure 1. Curve a shows the undeflected beam with the analyzer magnet turned off. Curve b shows the spreading of the beam with a low magnetic field. Gradient is not great enough to cause separation of components. Curve c shows separation in high field gradient. Notice the range of deflections, due principally to spread of velocities in beam from oven.

along the z-direction,* and its magnitude decreases with increasing z. For any orientation of the standard device, the field strength decreases as one moves in the direction along which $\vec{B}$ points.

The behavior of an atomic beam passing through an analyzer depends on the nature of the atoms in the beam. Of the atoms in the periodic table, approximately 12 percent--in their ground state--undergo no deflection whatever. Our interest centers on the 90 or so percent that are deflected, on excited atoms that are deflected, as well as the neutrons and molecules that we have subsumed in this chapter under the generic name "atomic beams." For beams of these particles, passage through the analyzer leads to a splitting into two, three, or more beams; the number of beams depends on the type of "atom" and its condition of excitation. The remarkable fact that this splitting is always into a set of discrete beams allows us to apply the concept quantum state (actually "spin state") to these atoms and to describe their behavior using mathematical machinery adapted from the description of photon polarization.

Figure 4 shows the idealized operation of a magnetic analyzer for the case of a beam of atoms that splits into two beams. Figure 5 shows an actual intensity profile obtained by using the apparatus of Figure 1. The beam profiles are obtained by moving a hot-wire detector across the beams. The deflections of the two components are equal in magnitude and opposite in sense.

By using a second analyzer oriented identically to the first, we could verify in the usual manner that the particles in the beams that emerge from the analyzer are in a set of orthogonal quantum states. Each beam does not split again when it passes through the second analyzer, but is merely deflected further in the same direction. Figure 6 shows idealized experiments for the case of a two-state system. The behavior is analogous to that of photon beams when they pass through successive identically oriented calcite analyzers. By measuring the intensities of input and output beams, one can verify also that the set of states defined by the analyzer is complete: for any input, the intensity of the input is equal to the sum of the intensities of the output beams.

In the case of the polarization of electromagnetic radiation, the action of a calcite crystal in splitting an incident beam into two can be understood classically. This is not so for the splitting of an atomic beam first detected by Stern and Gerlach. One can predict on the basis of a simple classical model that an atom should be deflected when it passes through a region of inhomogeneous magnetic field. But the separation into distinct beams cannot be explained classically.

The classical model results from assuming that each atom behaves like a tiny magnetic

---

* The fact that the field is inhomogeneous necessarily implies that it has a non-zero component in the direction transverse to the N-S axis. A non-uniform field that points in a single direction throughout a small region would violate the condition $\text{div } \vec{B} = 0$ required by Maxwell's equations.

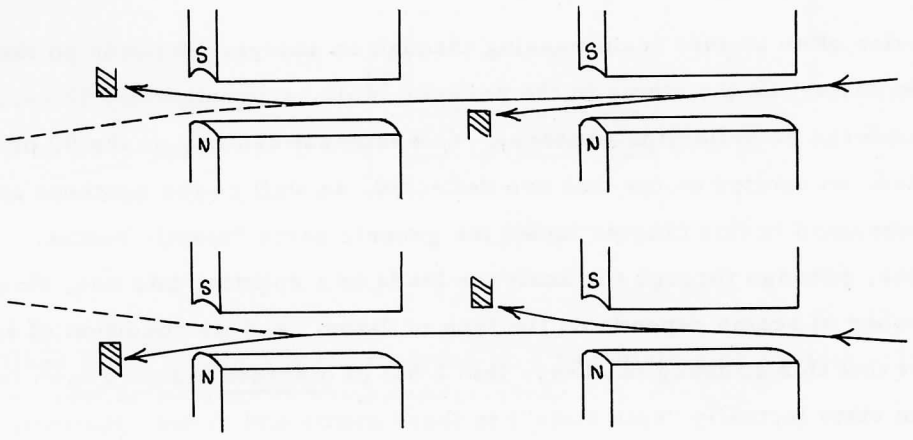
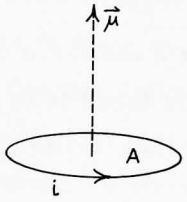


Figure 6. Idealized experiments used to verify that the output beams of the Stern-Gerlach apparatus represents orthogonal quantum states.

dipole. This assumption seems a plausible one, since the moving charges within an atom can be considered as current loops; and according to electromagnetic theory a small current loop carries a magnetic dipole moment. Box one develops the classical model of the atom as a magnetic dipole and presents the resulting classical predictions. Briefly, a dipole in a uniform field experiences no net force, but only a torque that causes it to precess like a top about the field direction. During this precession the angular momentum vector of the atom traces out a cone about the field direction. But in the inhomogeneous magnetic field shown in Figure 3 the atom experiences, in addition, a net force proportional to $\cos\theta$, where θ is the angle between the magnetic moment and the field direction. This force causes a transverse deflection of the atom, a deflection that is likewise proportional to $\cos\theta$. According to this classical model, therefore, each individual atom ought to be deflected either upward or downward with greater or lesser deflection (or none at all) according to its value of μ_z (Equation 3). The source of the cesium beam in Figure 1 is an oven with a small aperture. Presumably atoms emerge from such a source with their angular momenta and magnetic moments randomly oriented. The corresponding values of μ_z form a continuous band between the value $-\mu$ and the value $+\mu$. The classical model therefore predicts that the inhomogeneous field should spread the beam in a continuous distribution transverse to its direction of motion.

Box 1. Magnetic Moment of the Atom--Classical Model

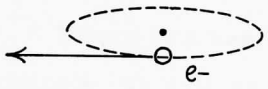


The magnetic moment of a classical current loop has the magnitude

$$\mu = iA/c$$

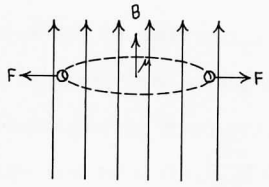
and the direction given by the "right hand rule". (Reference: Purcell, page 364.)

For a classical model of the atom, one argues that an electron moving in a closed orbit with velocity v constitutes an effective positive current of magnitude



$$i = -\nu e = -ev/2\pi r$$

in a direction opposite to the motion of the negatively charged electron. The area A of a circular orbit is



$$A = \pi r^2$$

Therefore the magnitude of the dipole moment is

$$\mu = -(ev/2\pi r)\pi r^2/c = \frac{-evr}{2c}$$

which can be written in the form

$$\mu = -eL/(2mc)$$

where $L = mvr$ is the angular momentum of the electron in its orbit. The proportionality between magnetic moment and angular momentum plays an important part in the theory. It is convenient to measure L in quantum units of angular momentum $\hbar = h/2\pi$

$$L = \hbar\ell$$

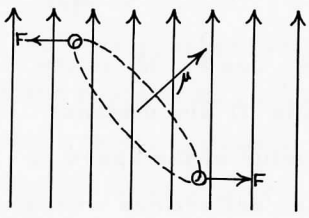
Then

$$\mu = -(e\hbar/2mc)\ell \qquad (1)$$

where the minus sign comes from the negative charge of the electron. The constant $(e\hbar/2mc)$ is the natural unit of atomic magnetic moment, and is called the Bohr magneton μ_B. To make the model slightly more general, we may assume that μ and ℓ are proportional, but the constant of proportionality is not necessary given by μ_B. Introduce a factor g ("Lande g-factor") to be determined by experiment for each type of atom.

$$\mu = \mu_B g \ell \qquad (2)$$

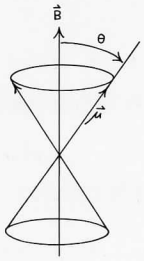
All of our ignorance is then summarized in the value of g.



In a uniform magnetic field parallel to μ, a current loop experiences a stress that tends to stretch it, but no net force. If the magnetic moment is inclined to the field, there is still no net force, but there is a torque that tends to turn the dipole moment into the field direction. The torque M may be shown to have the value

$$\vec{M} = \vec{\mu} \times \vec{B}$$

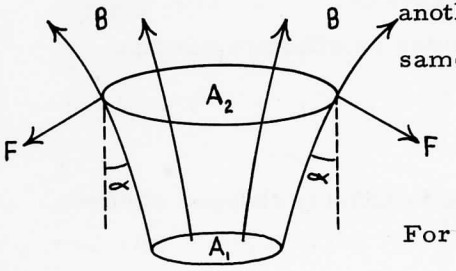
If the magnetic dipole represents a circulating electron, as in the classical model under consideration, the torque causes a precession



of the angular momentum, similar to the behavior of a gyroscope. The angular momentum traces out a cone whose axis lies along the field direction. The angular precession velocity is

$$\omega_p = \mu B/L = (g\mu_B/\hbar)B$$

In a non-uniform magnetic field the loop experiences a net force of magnitude $F_z = F \sin \alpha$. One way of calculating this force is to draw another circle below the orbit in such a way that the flux Φ is the same through both circles. Then

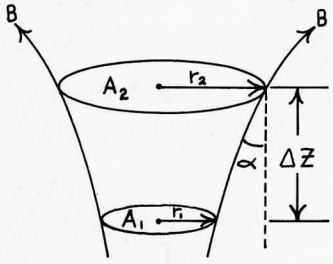


$$\Phi \approx B_1 A_1 = B_2 A_2$$

$$B_1 \pi r_1^2 = B_2 \pi r_2^2$$

For small angle α

$$r_2 = r_1 + \Delta z \tan \alpha \approx r_1 + \Delta z \sin \alpha$$



Neglect terms quadratic in Δz

$$B_1 r_1^2 \approx B_2 (r_1^2 + 2 r_1 \Delta z \sin \alpha)$$

Whence

$$\sin \alpha \approx -\frac{r_1^2}{2B_1} \frac{B_2 - B_1}{\Delta z} \approx -\frac{r}{2B} \frac{\partial B}{\partial z}$$

$$F_z = F \sin \alpha = \left(\frac{evB}{c}\right) \cdot \left(-\frac{r}{2B} \frac{\partial B}{\partial z}\right)$$

which can be written as

$$F_z = -(e/2mc)(mvr) \frac{\partial B}{\partial z} = -(eL/2mc) \frac{\partial B}{\partial z} = -(e\hbar/2mc) \ell \frac{\partial B}{\partial z}$$

Using Eq. 1 we can write the force exerted on the loop when $\vec{\mu}$ is parallel to $\vec{B}$

$$F_z = \mu \frac{\partial B}{\partial z}$$

From the value of this force we can calculate approximately the deflection d of the beam at the detector.

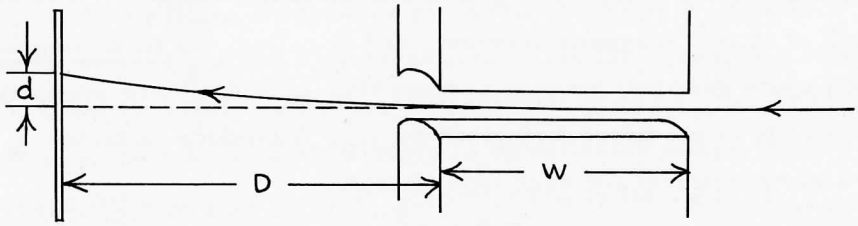


Figure 7b. Dimensions used to derive classical expression for beam deflection.

$$d = (WD\mu_z \frac{\partial B}{\partial z})/(u^2 M) \qquad \text{(for } D \gg W\text{)} \qquad (3)$$

Here μ_z is the z-component of the magnetic moment of the atom: $\mu_z = \mu \cos \theta$, M is the mass of the atom, u its speed, W the length of the deflecting magnet, and D the distance from magnet to detector. Notice the sensitivity of this deflection to the value of the speed u of the atom. Equation 3 predicts that for random orientations of atoms the deflections will take on a continuous range of values. This is in marked contrast to the splitting into discrete beams observed in Experiment (see curve c in Figure 5).

The classical prediction is contradicted by the result of experiment: curve c of Figure 5 shows a cesium beam split into two distinct components. In terms of the classical model, this result suggests that the atomic dipole moments are not oriented at random. Instead, they appear to form two groups, the members of each group having (very nearly) the same value of μ_z. The two "allowed" values of μ_z have equal magnitude and opposite sign. If we were to take this classical picture literally, we would conclude that the magnetic moment vectors lie along two cones, as shown in Figure 7a in the box. The opening angle of these cones is $\theta = \cos(\mu_z/\mu)$. (The value of μ_z can be inferred from the magnitude of the beam deflections, but without knowing the magnitude of μ we could not determine θ.)

There is a serious difficulty with the classical picture. The oven from which the atoms originate has nothing to do with the magnetic field of the analyzer. Yet, it is claimed, this oven emits atoms whose magnetic moments lie on cones whose axes are along the field direction. If we were to carry out the same experiment with a rotated analyzer, the beam would split exactly as before, but along the new field direction of the analyzer. We would have to conclude that the magnetic moments of the atoms from the same oven now lie on a different pair of cones that have the new field direction as axis. Clearly the "quantization axis" is established by the analyzer and not the source; a correct description of the splitting must reflect this fact. The difficulty here mentioned arises from our insistence on assigning a definite direction to the magnetic moment of each atom. As we shall see shortly, a modification of this point of view leads to the resolution of the problem.

As we pointed out at the beginning of this section, the splitting of an atomic beam into cleanly separated components, each of which travels in a unique direction, is an idealization impossible to achieve in practice. The principal cause of difficulty is the fact that sources generally provide beams in which the atoms have a range of velocities. When such a beam passes through the analyzer the slower atoms spend more time in the inhomogeneous field and are therefore deflected more than are the fast atoms (Eq. 3 in Box 1). As a result, even with a two-state system, which most closely approximates the idealized behavior, the directions of the atoms within each of the separated beams exhibit considerable spread, as observed in curve c of Figure 5. For beams that would ideally be split into three or more components, the spread of velocities causes considerable overlap among the various components, and may even obscure the pattern entirely. Figure 8 shows a typical example of this effect with a four-state system. If we were presented with the intensity profile shown in this figure, we surely could not decide on the basis of this evidence alone that the pattern represents a four-state rather than a two-state system. When an atom has more than two spin states, the number of states is generally inferred from other types of experiments (for example the resonance experiments to be discussed in a later chapter). There is every reason to believe that a sufficiently monoenergetic beam of such atoms could be separated by an inhomogeneous-field magnetic analyzer. We shall

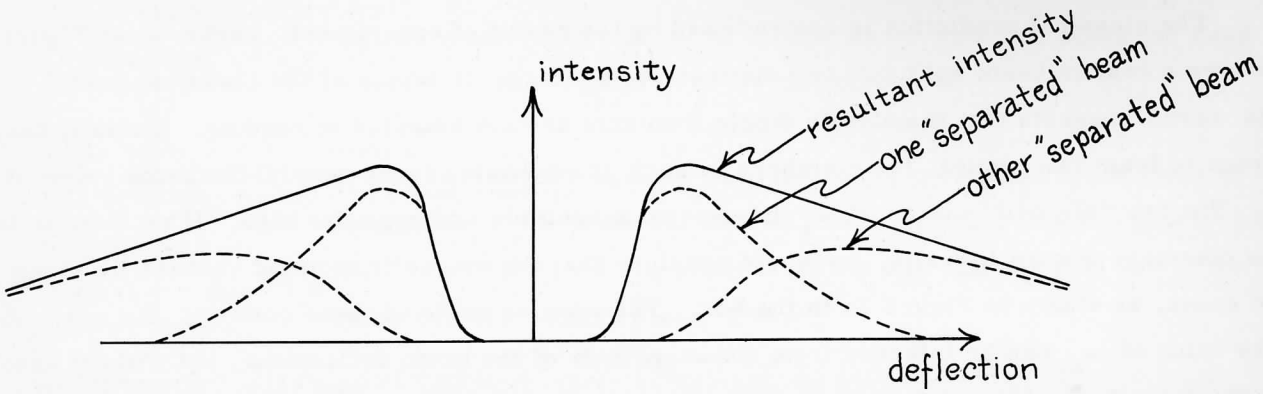


Figure 8. Velocity spread destroys clean separation of four beams, as illustrated in this theoretical curve. (Molecular Rays Ronald G. J. Fraser, New York, Macmillan Company, Cambridge, England at the University Press, 1931.)

therefore invoke in our discussion ideal analyzers for systems with any number of states, even though a clean separation has actually been achieved only for the two-state system.

4. Two-state probability amplitudes

We now investigate the properties of quantum states defined by a magnetic analyzer, concentrating on the simplest case--a system that is split into only two components. Specific examples of such systems are beams of cesium atoms and neutrons. These beams behave in identical fashion in the experiments we shall describe, except that the magnitudes of the deflections of the outgoing beams differ. We shall interpret these differences as being due to differences in the magnitudes of the dipole moments of the particles involved.

The atoms that emerge in the upper output beam of the analyzer of Figure 2 (that is, the atoms that have been deflected in the direction of the field) all have the same negative value of μ_z the atoms in the other output beam have an equal but positive μ_z. We could use this identification to classify the states. However, because the angular momentum is a more fundamental quantity in the description of atomic states than the magnetic moment, it is preferable to classify the states according to their angular momentum rather than according to their magnetic moment. The two quantities are very closely related; in the classical model, as we showed in Box 1, they are directly proportional to one another. Without insisting on a strict proportionality, we observe that, because atomic magnetic moments are principally due to electrons which are negatively charged, an atom with a positive component of magnetic moment in a given direction is likely to have a negative component of angular momentum, and vice versa. This expectation will later be confirmed. Accordingly, we define as $|+z\rangle$ the state of the atoms deflected into the upper beam by the standard analyzer of Figure 4, and $|-z\rangle$ the state of the atoms deflected into the lower beam. The labels + and - then refer to the sign of the component of angular momentum along the direction defined by the analyzer.

The states $|+z\rangle$ and $|-z\rangle$ are the basis states for the spin system. They are commonly called "spin up" and "spin down" respectively. Use of this language does not imply that when an

atom is in one of these states its angular momentum points directly along or directly opposite to the field direction; as we have already remarked and shall see again below, such a strict adherence to the classical model cannot be justified.

Any other orientation of the analyzer, say z', similarly defines a pair of basis states which we can call $|+z'\rangle$ and $|-z'\rangle$; these describe the atoms deflected respectively along and opposite to the z' direction by a z' analyzer. An atom in the state $|+z'\rangle$ has a positive component of spin angular momentum in the z' direction, and a negative component of magnetic moment.

Next consider the projection probability between two states belonging to different bases, say state +z and state +z'. The experiment that measures these probabilities is sketched in Figure 9. For concreteness we take the beam direction to be the y axis; then z' is in the xz plane and makes an angle θ with the z axis.

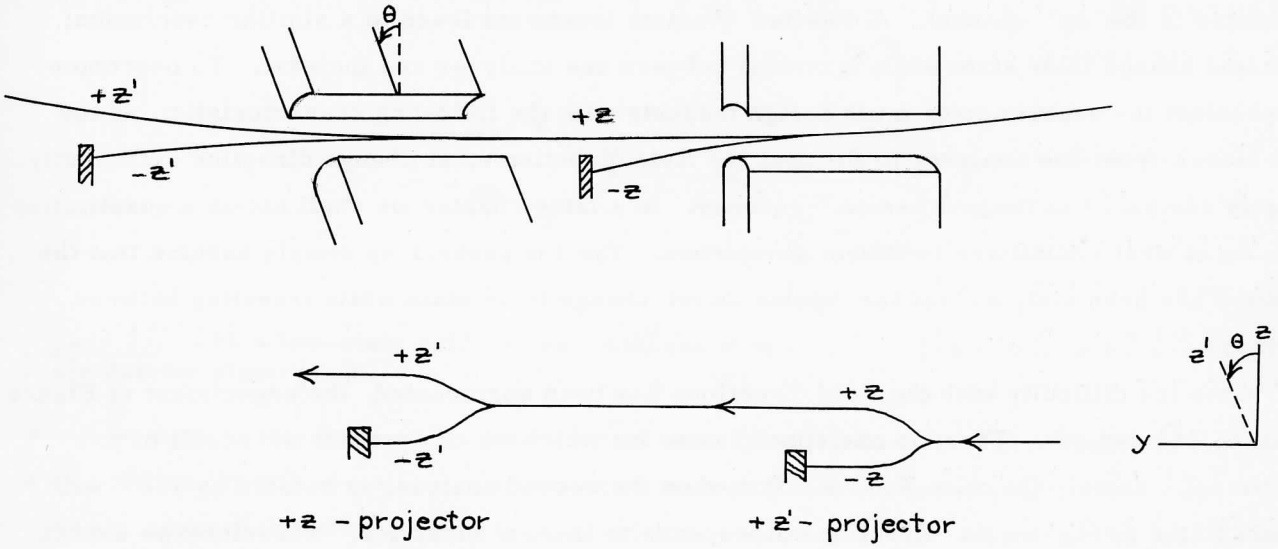


Figure 9. Experiment to measure the quantity $|\langle +z' | +z \rangle|^2$

There is a serious technical problem associated with all experiments of this type. The diagram implies that the atoms pass from a region where the magnetic field points in the z direction, into one in which the field points in the z' direction. A completely sudden and discontinuous change of direction for $\vec{B}$ would violate the laws of magnetostatics; a field with zero curl must change direction continuously. Unless special precautions are taken, the field configuration along the path of the beams will change gradually, as indicated schematically in Figure 10. Inside each magnet the field is strong and points in the designated direction. Just outside each magnet, the field is weaker but its direction is still correlated with the direction of the field in the interior. In the space between the magnets, the field is so weak that its direction is hard to determine. But it is clear that the change in field direction from z to z' takes place gradually. Under such conditions the experiment of Figure 9 will not measure what it is

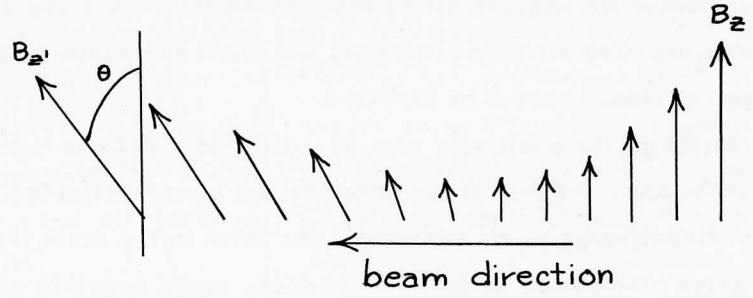


Figure 10. Gradual change of direction of magnetic field B between two analyzers z and z'.

intended to measure. In terms of the classical picture we may say that the dipoles, which leave the first analyzer pointing along the field, "follow the field" as it changes direction. When they enter the second analyzer, they point along the direction of the new field and will all be counted in the $+z'$ channel. A detailed quantum treatment leads to a similar conclusion: the atoms change their state while traveling between one analyzer and the next. To overcome this problem the experimenter must design magnets with the following characteristic: as the atom moves from one analyzer to the next the field direction must change direction sufficiently abruptly compared to the precession frequency. In a later chapter we shall obtain a quantitative estimate of what constitutes sufficient abruptness. For the present we simply assume that the criterion has been met, so that the dipoles do not change their state while traveling between analyzers.

Once the difficulty with the field directions has been surmounted, the experiment of Figure 9 can be carried out. There is one special case for which we can predict the result of this experiment, namely the case $\theta = \pi$. For when the second analyzer is rotated by 180° with respect to the first, the $+z$ direction corresponds to the $-z'$ direction. Therefore we expect that the atoms that emerge from the first analyzer in the state $|+z\rangle$ should all be deflected into the $-z'$ channel of the second analyzer. (Notice that this again corresponds to an upward deflection of the beam.) This prediction is readily confirmed by experiment; the projection probabilities are

$$\left.\begin{aligned} |\langle -z'|+z\rangle|^2 &= 1 \\ |\langle +z'|+z\rangle|^2 &= 0 \end{aligned}\right\} \quad \theta = \pi \tag{4}$$

Similarly one finds

$$\left.\begin{aligned} |\langle -z'|-z\rangle|^2 &= 0 \\ |\langle +z'|-z\rangle|^2 &= 1 \end{aligned}\right\} \quad \theta = \pi \tag{5}$$

In the general case, there is no way for us to predict the outcome of the experiment. The results turn out to be as follows:

$$\begin{aligned} |\langle +z'|+z\rangle|^2 &= |\langle -z'|-z\rangle|^2 = \cos^2\frac{\theta}{2} \\ |\langle +z'|-z\rangle|^2 &= |\langle -z'|+z\rangle|^2 = \sin^2\frac{\theta}{2} \end{aligned} \tag{6}$$

the appearance of the half-angle in these formulas will have important consequences.

These results indicate anew the limitations of the classical model in which the magnetic moment and angular momentum of an atom are supposed to point in a specific direction. Consider for concreteness the case $\theta = \pi/2$: the magnetic field in the second analyzer of Figure 11 points along the positive x direction. Eq. (6) for this case states that

$$|\langle +x|+z\rangle|^2 = |\langle -x|+z\rangle|^2 = \frac{1}{2}$$

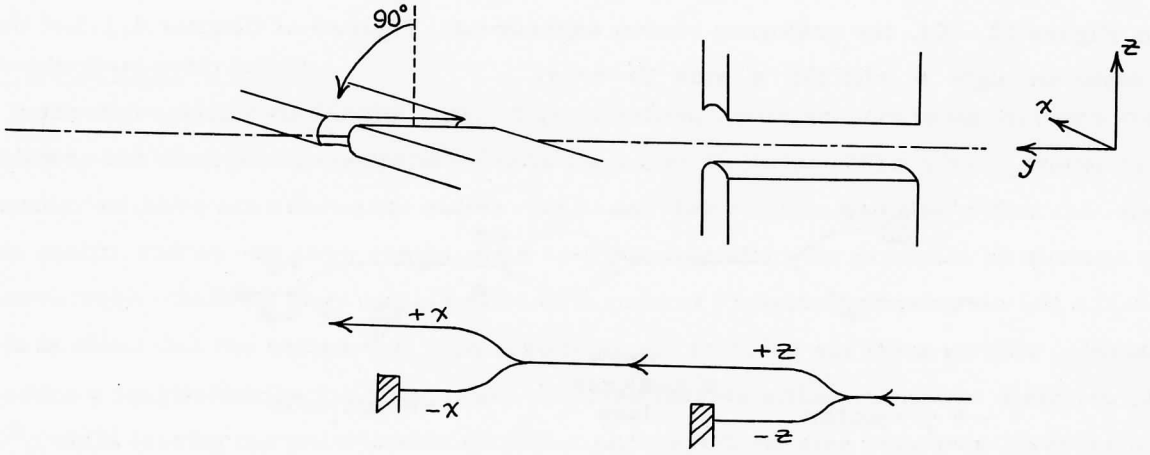


Figure 11. For a two-state spin system, half the atoms in state $|+z\rangle$ will be passed by a projector for state $|+x\rangle$.

When an atom in the state +z is tested by an x-analyzer, it has equal probability of emerging with positive or with negative x-component of magnetic moment (or angular momentum). Moreover the separation between the +x beam and the -x beam in the x-analyzer is just as great as the separation between the two beams in the z-analyzer from which the +z beam was obtained. How can a dipole "pointing in the +z direction" have two discrete non-zero components "pointing in opposite directions along the x axis?" The results of real experiments with atomic beams cannot be reconciled with the classical picture of the atom presented in Box 1, according to which an atomic magnetic moment "points" in one direction or another. Instead one thinks correctly in terms of basis states determined by an analyzer with a particular orientation. One of the beams from a first analyzer can be analyzed in terms of a different basis by passing the beam through a second analyzer at the new orientation. The experiment shows that the quantum state $|+z\rangle$ is a linear superposition of the states $|\pm x\rangle$.

Having determined the projection probabilities for the two-state system we next inquire about the values of the corresponding projection amplitudes. We have of course already anticipated that such amplitudes exist, by writing the probabilities in the standard bracket notation. In our earlier discussion of photon states we demonstrated directly that amplitudes are necessary, by considering experiments with analyzer loops. Unfortunately, an analyzer loop for atomic

beams has yet to be constructed; the technical problems associated with such a project are severe. But the results of many other experiments support the hypothesis that atomic states are described by probability amplitudes that can interfere, and whose absolute squares determine probabilities. The success of the theory based on this hypothesis allows us to employ the analyzer loop for atomic states as a conceptual device in thought-experiments, even though no such device actually exists. (The magnetic analyzer for a many-state system will play a similar role in our later considerations.)

The thought-experiment that provides information concerning phases of amplitudes is shown in Figure 12. (Cf. the analogous photon experiment, Figure 6 of Chapter 4.) Let the z' axis make an angle θ with the z axis ($0<\theta<\pi$).

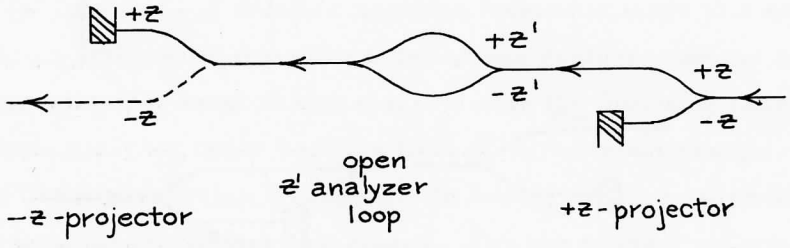


Figure 12. Use of idealized analyzer loop to determine phases of projection amplitudes.

Since as usual, the open analyzer loop has no effect on the state of the particles that pass through it, no particles are counted in the -z channel of the second analyzer of Figure 12. Hence we can write

$$\langle -z|+z\rangle = \langle -z|+z'\rangle\langle +z'|+z\rangle + \langle -z|-z'\rangle\langle -z'|+z\rangle = 0 \qquad (7)$$

and similarly,

$$\langle +z|-z\rangle = \langle +z|+z'\rangle\langle +z'|-z\rangle + \langle +z|-z'\rangle\langle -z'|-z\rangle = 0 \qquad (8)$$

Eqs. 4 and 5 tell us the magnitudes of all the amplitudes that appear in (7) and (8). Each of the latter equations provides one condition among four phase factors. Just as in the case of photon states, the absolute phases are a matter of convention; the general phase convention to be described in the next section assigns to the amplitudes in equations (7) and (8) the following values:

$$\begin{aligned} \langle +z|+z'\rangle &= \langle +z'|+z\rangle = \cos\frac{\theta}{2} \\ \langle -z|+z'\rangle &= \langle +z'|-z\rangle = \sin\frac{\theta}{2} \\ \langle +z|-z'\rangle &= \langle -z'|+z\rangle = -\sin\frac{\theta}{2} \\ \langle -z|-z'\rangle &= \langle -z'|-z\rangle = \cos\frac{\theta}{2} \end{aligned} \qquad (9)$$

The corresponding expressions for the kets $|+z'\rangle$ and $|-z'\rangle$ are

$$|+z'\rangle = |+z\rangle \cos\frac{\theta}{2} + |-z\rangle \sin\frac{\theta}{2}$$
$$|-z'\rangle = -|+z\rangle \sin\frac{\theta}{2} + |-z\rangle \cos\frac{\theta}{2} \qquad (10)$$

The kets $|+z'\rangle$ and $|-z'\rangle$ are orthogonal, as they ought to be. For $\theta = \pi/2$ we have in particular

$$|+x\rangle = (|+z\rangle + |-z\rangle)1/\sqrt{2}$$
$$|-x\rangle = (-|+z\rangle + |-z\rangle)1/\sqrt{2} \qquad (11)$$

Similar expressions can be written for the dual vectors $\langle \pm z'|$, etc.

5. Longitudinal polarization

In the discussion thus far the "axis of quantization," with respect to which basis states are defined, has always been perpendicular to the beam direction. For a beam traveling in the y direction, we have considered the states $|\pm z\rangle$ and $|\pm x\rangle$. One can also define the state $|+y\rangle$ for this beam; that is, an atom can have its spin pointing along its direction of motion, as well as transversely. Such an atom is said to be in a state of "longitudinal polarization."

It is clear that the beams that emerge from an analyzer are transversely polarized. To produce a longitudinally polarized beam requires that we either rotate the direction of motion by 90°, while leaving the polarization direction unchanged, or else rotate the polarization while leaving the direction of motion unchanged. The first possibility is hard to accomplish. In principle, one could use a gravitational force to turn the beam; gravity presumably would turn the beam without affecting the orientation of the spin. However, gravitational forces are so weak that no one would seriously consider building a gravity-based beam turner. On the other hand, a rotation of the magnetic moment is readily accomplished by passing the beam through a uniform magnetic field, as shown schematically in Figure 13. With the $-z$ channel of the analyzer blocked, the emerging beam is in the state $|+z\rangle$ (spin up, magnetic moment down). When this beam passes through a uniform magnetic field in the x direction its motion is unaffected, but the magnetic moment vector precesses around the x axis. This result was derived from the

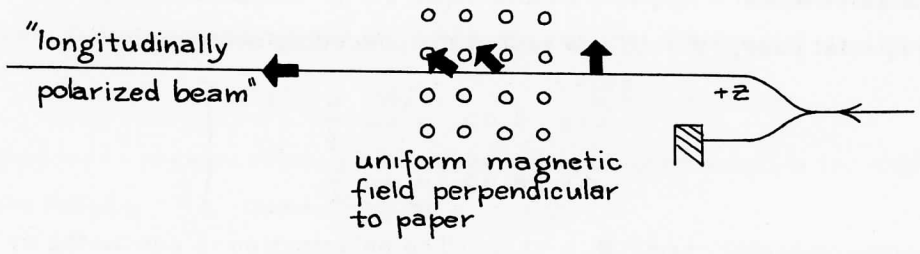


Figure 13. Use of uniform magnetic field to rotate spin direction of atoms.

classical model in Box 1. If the strength of the field and the length of the field region bear the proper relation to the particle velocity, the total precession can be made to be 90° and the at atoms emerge polarized longitudinally. Of course the precession argument is based on the classical model, whose difficulties have already been noted. The conclusion must be tested experimentally. An experiment described in the next chapter verifies that the classical precession argument leads to correct conclusions in this case.

Atomic beams can be polarized transversely to their direction of motion or along their direction of motion--or indeed at any angle θ, ϕ (Figure 14, in which the beam direction is along the y axis). Polarization of atoms at an arbitrary angle can be accomplished by a variation of the procedure of Figure 13: Choose the magnitude and direction of a uniform magnetic field so that the initial $+z$ atoms precess to the desired final state. Detection of atoms in the state specified by θ, ϕ can be carried out by reversing this procedure: Find a uniform magnetic field such that after passage through it, all atoms are transmitted by a $+z$-projector.

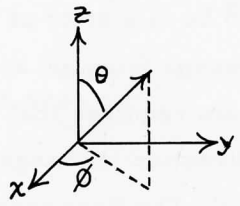


Figure 14. Designation of an arbitrary direction using the angles θ, ϕ.

6. The state vector for the two-state system

Having established the possibility of longitudinal polarization for atomic beams, we are in a position to investigate the general state of a two-state system. Let $|\theta, \phi\rangle$ denote the state polarized in the direction defined by the polar angles θ, ϕ. This ket describes the polarization state of the beam that emerges from the + channel of an analyzer whose axis points in the (θ, ϕ) direction. By turning the beam (gravity) or precessing the spin (uniform magnetic field), we can assign the polarization state $|\theta, \phi\rangle$ to a beam moving in any direction. We now ask, what are the amplitudes $\langle +z \mid \theta, \phi\rangle$ and $\langle -z \mid \theta, \phi\rangle$? We already know the answer to this question for one special case, $\phi = 0$; we saw in the preceding section that

$$\left.\begin{aligned} \langle +z \mid \theta, 0\rangle &= \cos\frac{\theta}{2} \\ \langle -z \mid \theta, 0\rangle &= \sin\frac{\theta}{2} \end{aligned}\right\} \qquad (12)$$

Consider now another special case: $\theta = \pi/2$. The polarization is now in the xy plane. The amplitudes in this case both have the magnitude $1/\sqrt{2}$. Within an overall phase they can be written in the form

$$\left.\begin{aligned} \langle +z \,|\, \tfrac{\pi}{2}, \phi \rangle &= \frac{1}{\sqrt{2}}\, e^{i\alpha(\phi)} \\ \langle -z \,|\, \tfrac{\pi}{2}, \phi \rangle &= \frac{1}{\sqrt{2}}\, e^{-i\alpha(\phi)} \end{aligned}\right\} \qquad (13)$$

where α is some unknown function of ϕ.

To determine the dependence of α on ϕ we consider the amplitude $\langle \frac{\pi}{2}, \phi_2 | \frac{\pi}{2}, \phi_1 \rangle$ between two states both of the form (13). The absolute square of this amplitude is the projection probability for a beam polarized along the direction $(\pi/2, \phi_1)$, when the projector axis points along the direction $(\pi/2, \phi_2)$. According to the basic experiment described in the last section, which determines projection probabilities, we have

$$\left| \langle \pi/2, \phi_2 \,|\, \pi/2, \phi_1 \rangle \right|^2 = \cos^2 \left[\frac{\text{angle between directions } (\pi/2, \phi_2) \text{ and } (\pi/2, \phi_1)}{2} \right]$$

$$= \cos^2 \left(\frac{\phi_2 - \phi_1}{2} \right) \qquad (14)$$

See Fig. 14.

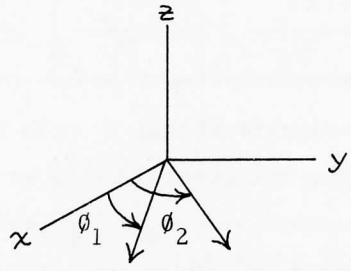


Figure 14. Special case: beam polarized at $(\frac{\pi}{2}, \phi_1)$, projector oriented at $(\frac{\pi}{2}, \phi_2)$

Expanding the amplitudes $\langle \frac{\pi}{2}, \phi_2 | \frac{\pi}{2}, \phi_1 \rangle$ in terms of the $\pm z$ basis and using equations 13, we obtain

$$\langle \tfrac{\pi}{2}, \phi_2 | \tfrac{\pi}{2}, \phi_1 \rangle = \langle \tfrac{\pi}{2}, \phi_2 \,|\, +z \rangle \langle +z | \tfrac{\pi}{2}, \phi_1 \rangle + \langle \tfrac{\pi}{2}, \phi_2 | -z \rangle \langle -z | \tfrac{\pi}{2}, \phi_1 \rangle$$

$$= \frac{1}{2} \left[e^{i\left(\alpha(\phi_1) - \alpha(\phi_2)\right)} + e^{-i\left(\alpha(\phi_1) - \alpha(\phi_2)\right)} \right]$$

$$= \cos \left[\alpha(\phi_1) - \alpha(\phi_2) \right] \qquad (15)$$

We have assumed that $\langle A \,|\, B \rangle^* = \langle B \,|\, A \rangle$ is true for spin states as well as for photon states. Comparing equations 15 and 14 we see that the possible solutions are

$$\alpha(\phi) = \pm\phi/2 \qquad \text{or} \qquad \pm\left(\frac{\phi}{2} + \text{constant}\right) \qquad (16)$$

The choice that leads to expressions which agree with standard notation is $\alpha(\phi) = -\phi/2$. With this choice the ket $|\frac{\pi}{2}, \phi\rangle$ takes the form

$$|\tfrac{\pi}{2}, \phi\rangle = |+z\rangle \left(\frac{1}{\sqrt{2}}\, e^{-i\phi/2} \right) + |-z\rangle \left(\frac{1}{\sqrt{2}}\, e^{i\phi/2} \right) \qquad (17)$$

Finally, we turn to the general case in which both θ and ϕ are arbitrary. The absolute squares of the amplitudes are given by equations (6). Defining phase factors in a manner analo-

gous to (13), we write

$$\left.\begin{aligned} \langle +z \mid \theta, \phi \rangle &= \cos\frac{\theta}{2}\, e^{i\alpha(\theta,\phi)} \\ \langle -z \mid \theta, \phi \rangle &= \sin\frac{\theta}{2}\, e^{-i\alpha(\theta,\phi)} \end{aligned}\right\} \qquad (18)$$

The phase factor α might of course depend on θ as well as on ϕ; all we have learned thus far is that in the special case $\theta = \pi/2$, α can be chosen to be $-\phi/2$. We proceed just as in that special case, requiring that the projection probability between two arbitrary states be equal to $\cos^2$ of half the angle between the corresponding directions, that is, we require:

$$\left| \langle \theta_2, \phi_2 \mid \theta_1, \phi_1 \rangle \right|^2 = \cos^2 \frac{\gamma}{2}, \qquad (19)$$

where γ is the angle between the direction (θ_1, ϕ_1) and (θ_2, ϕ_2). (Fig. 15.)

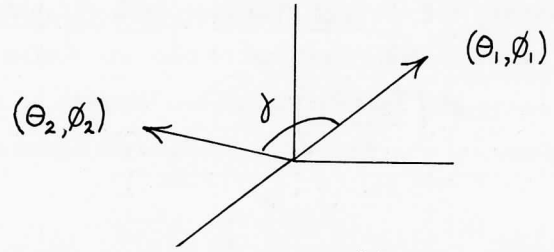


Figure 15. General case: Beam polarized at (θ_1, ϕ_1), projector oriented at (θ_2, ϕ_2).

Complete analysis shows (see the exercises) that α is in fact independent of θ, and can be chosen to be $-\phi/2$ in general. Thus the general ket is written, in our convention, as

$$|\theta, \phi\rangle = |+z'\rangle = |+z\rangle \cos\frac{\theta}{2}\, e^{-i\phi/2} + |-z\rangle \sin\frac{\theta}{2}\, e^{i\phi/2} \qquad (20)$$

The angles θ, ϕ specify an arbitrary direction in space. We have placed the z' axis along this arbitrary direction and have called this state $|+z'\rangle$. A state orthogonal to $|\theta, \phi\rangle$ can then be called $|-z'\rangle$. One way to write this orthogonal ket is

$$|-z'\rangle = -|+z\rangle \sin\frac{\theta}{2}\, e^{-i\phi/2} + |-z\rangle \cos\frac{\theta}{2}\, e^{i\phi/2} \qquad (21)$$

Any state orthogonal to (20) must be described by (21) times an overall phase factor. The particular phase in (21) is chosen to be consistent with a slightly more general (but algebraically more complicated) treatment in which one is concerned also with the directions of the x' and y' axes.*

The states $|\pm x\rangle$ have already been written down in Eq. 11. The direction +y corresponds to $\theta = \pi/2$, $\phi = \pi/2$. Therefore, according to (20)

$$|+y\rangle = |\pi/2, \pi/2\rangle = \left(|+z\rangle + |-z\rangle\, i\right) e^{-i\pi/4}/\sqrt{2} \qquad (22)$$

According to (21) the state orthogonal to $|+y\rangle$ is defined as

$$|-y\rangle = \left(-|+z\rangle + |-z\rangle\, i\right) e^{-i\pi/4}/\sqrt{2} \qquad (23)$$

Using equations (11), (22), and (23) we can construct the following table of projection amplitudes.

*The ket $|-z'\rangle$ must be polarized in the direction opposite to (θ, ϕ). This opposite direction has the angles $\pi-\theta$, $\phi+\pi$. If one substitutes these values into (20), one indeed obtains (21) but multiplied by an overall phase factor $-i$. This additional phase factor does not appear in the full treatment that includes the orientation of the x' and y' axes.

Table 3. Some projection amplitudes for a two-state spin system.

TO STATE \ FROM STATE	$\|+x\rangle$	$\|-x\rangle$	$\|+y\rangle$	$\|-y\rangle$	$\|+z\rangle$	$\|-z\rangle$
$\langle+x\|$	1	0	$1/\sqrt{2}$	$i/\sqrt{2}$	$1/\sqrt{2}$	$1/\sqrt{2}$
$\langle-x\|$	0	1	$1/\sqrt{2}$	$1/\sqrt{2}$	$-1/\sqrt{2}$	$1/\sqrt{2}$
$\langle+y\|$	$1/\sqrt{2}$	$-i/\sqrt{2}$	1	0	$\frac{e^{i\pi/4}}{\sqrt{2}}$	$\frac{e^{-i\pi/4}}{\sqrt{2}}$
$\langle-y\|$	$-i/\sqrt{2}$	$1/\sqrt{2}$	0	1	$\frac{-e^{i\pi/4}}{\sqrt{2}}$	$\frac{e^{-i\pi/4}}{\sqrt{2}}$
$\langle+z\|$	$1/\sqrt{2}$	$-1/\sqrt{2}$	$\frac{e^{-i\pi/4}}{\sqrt{2}}$	$\frac{-e^{-i\pi/4}}{\sqrt{2}}$	1	0
$\langle-z\|$	$1/\sqrt{2}$	$1/\sqrt{2}$	$\frac{e^{i\pi/4}}{\sqrt{2}}$	$\frac{e^{i\pi/4}}{\sqrt{2}}$	0	1

The properties of the spin states derived in this chapter can be written in the matrix notation developed in Chapter 6. In the z representation, an arbitrary spin state $|\psi\rangle$, characterized by its projection amplitudes $\langle+z|\psi\rangle$, $\langle-z|\psi\rangle$, can be written as a column matrix

$$\begin{pmatrix} \langle+z|\psi\rangle \\ \langle-z|\psi\rangle \end{pmatrix} \tag{24}$$

The corresponding dual vector $\langle\psi|$ is represented by the row matrix

$$\begin{pmatrix} \langle\psi|+z\rangle & \langle\psi|-z\rangle \end{pmatrix} = \begin{pmatrix} \langle+z|\psi\rangle^* & \langle-z|\psi\rangle^* \end{pmatrix} \tag{25}$$

The transformation to a different representation z' is accomplished by the transformation matrix

$$U = \begin{pmatrix} \langle+z'|+z\rangle & \langle+z'|-z\rangle \\ \langle-z'|+z\rangle & \langle-z'|-z\rangle \end{pmatrix} \tag{26}$$

Specifically we have (cf. Eq. 14 of Chapter 6)

$$\begin{pmatrix} \langle+z'|\psi\rangle \\ \langle-z'|\psi\rangle \end{pmatrix} = \begin{pmatrix} \cos\frac{\theta}{2}\, e^{i\phi/2} & \sin\frac{-\theta}{2}\, e^{i\phi/2} \\ -\sin\frac{\theta}{2}\, e^{i\phi/2} & \cos\frac{-\theta}{2}\, e^{i\phi/2} \end{pmatrix} \begin{pmatrix} \langle+z|\psi\rangle \\ \langle-z|\psi\rangle \end{pmatrix} \tag{27}$$

where θ and ϕ are the polar angles of the z' axis referred to the z axis.

7. Pauli matrices

Having defined basis kets for the two-state spin system and derived their transformation laws, we can proceed to discuss operators and their matrices in the same way as we did for photon states. An operator of particular interest is the one whose matrix in the z representation is

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{28}$$

This operator is called σ_z. The device corresponding to the operator σ_z is one that passes both $+z$ and $-z$ beams with no change of polarization, but with a phase factor which is $+1$ for the $+z$ state and -1 for the $-z$ state.

It is natural to associate with σ_z two additional operators σ_x and σ_y which fulfill analogous roles in the x- and y- representations. That is, we require that

$$\sigma_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{x-representation}$$
$$\sigma_y = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{y-representation} \tag{29}$$

The matrices of σ_x and σ_y in the z-representation are obtained from (29) by transforming in the usual manner. Specifically, one has (cf. section 3 of chapter 5):

$$(\sigma_x)_{z\ \text{rep}} = U(\sigma_x)_{x\ \text{rep}}\, U^{-1} \tag{30}$$

where the transformation matrix U is

$$U = \begin{pmatrix} \langle +z|+x\rangle & \langle +z|-x\rangle \\ \langle -z|+x\rangle & \langle -z|-x\rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{31}$$

The result is

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{z-representation} \tag{32}$$

Similarly, one finds

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \text{z-representation} \tag{33}$$

Thus far the subscripts x, y, z in the three σ operators merely designate the representation in which the corresponding matrix has the diagonal form (28). But one can go further and demonstrate that in a real sense the three operators behave like the components of a vector in ordinary three dimensional space. The identifying characteristic of a vector is the way its components transform under rotation of the coordinate axes. Let z' lie along the "direction of $\vec{\sigma}$." In analogy to the definition of σ_z above, we expect that the device described by $\vec{\sigma}$ will pass a beam in the state $+z'$ with a phase factor $+1$ and a beam in the state $-z'$ with a phase factor -1. The z'-component matrix of this device in the z' representation is therefore

$$\sigma_{z'} = \vec{\sigma} \cdot \hat{z'} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad (z'\text{-representation}) \tag{34}$$

Here $\hat{z}'$ is a unit vector pointing along the z' axis. Equation 34 can be transformed to the z-presentation (see the exercises). However, if $\vec{\sigma}$ behaves like a vector, then we ought to be able to predict the form of the z'-component in the z-representation from its components in that representation: σ_z, σ_x, and σ_y. Let θ and ϕ be the polar angles of z' with respect to the xyz axes (Figure 14). Then we have

$$\sigma_{z'} = \sigma \cdot \hat{z}' = \sigma_z \cos\theta + \sigma_x \sin\theta \cos\phi + \sigma_y \sin\theta \sin\phi \qquad \text{(z-representation)} \quad (35)$$

using equations (28), (32) and (33), (35) becomes

$$\cos\theta \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \sin\theta \left[\cos\phi \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \sin\phi \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \right] = \begin{pmatrix} \cos\theta & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta \end{pmatrix} \quad (36)$$

In the exercises it will be verified that matrix (34) has the form (36) when transformed to t the z-representation. Therefore we are justified in thinking of $\vec{\sigma}$ as a <u>vector operator</u>, that is, a vector in ordinary three-space each of whose components is an operator represented by a 2x2 matrix. These components transform under rotations just like the components of any vector.

The matrices σ_x, σ_y, and σ_z, known as Pauli spin matrices, play an important role in the theory of the two-state system. We shall see later on that the operator $1/2\,\vec{\sigma}$ represents the spin angular momentum. Here we merely list a few useful formal properties which may be verified from the explicit matrices in the z-representation.

(a) The square of each matrix is the unit matrix

$$\sigma_x \sigma_x = \sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (37)$$

(b) The matrices <u>anticommute</u>

$$\sigma_j \sigma_k = -\sigma_k \sigma_j \qquad j \neq k \quad (38)$$

(c)

$$\sigma_j \sigma_k = i\sigma_\ell \qquad \begin{matrix} j, k, \ell \text{ in cyclic order,} \\ \text{that is, xyz, yzx, or zxy} \end{matrix} \quad (39)$$

Chapter 8. Time-dependent State Vectors

1. Introduction

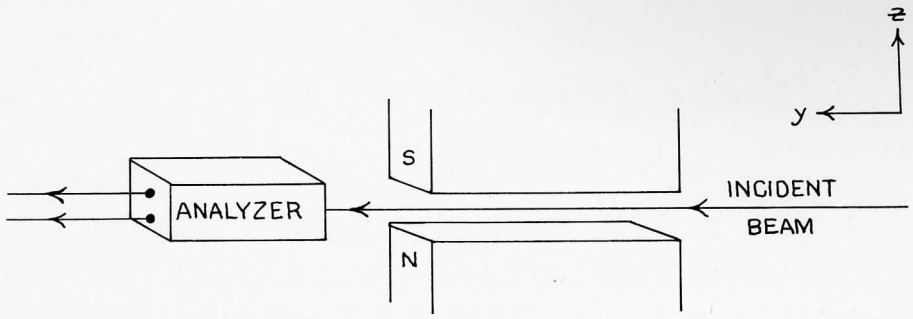
In this chapter we begin the study of the time dependence of amplitudes and state vectors by analyzing the behavior of a two-state spin system in a uniform magnetic field. We describe this behavior in terms of a state vector that changes with time. No matter what state the system is in initially, the time development of its state vector involves the energies of the set of basis states $|+z\rangle$ and $|-z\rangle$ associated with the magnetic field direction. The states $|+z\rangle$ and $|-z\rangle$ have another unique characteristic: any measurement on a system in either of these states yields a result that is independent of the time at which the measurement is carried out. For this reason the states $|+z\rangle$ and $|-z\rangle$ are called stationary states. Not all states of a two-state spin system in a magnetic field are stationary, but the time development of the state vectors of all states can be conveniently expressed in terms of the stationary basis set. Stationary states are so useful that their determination often constitutes the first step in studying a new physical system.

2. Two-state spin system in a uniform magnetic field

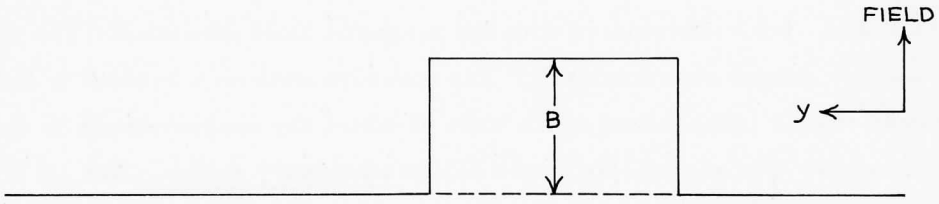
We want to investigate the time behavior of a two-state spin system in a uniform magnetic field of prescribed magnitude. To do this, we could place the atoms between the poles of a magnet, turn the field on suddenly at $t = 0$, and turn it off suddenly at some later time. Instead, it is more practicable to send a beam of atoms between the pole pieces (Figure 1a). Then the time an atom spends in the magnetic field can be controlled by controlling its speed. The condition that each atom experience a magnetic field that turns on and off suddenly in time then corresponds to the condition that the magnetic field rise and fall sharply in space. For the present, assume that this spatial change is in fact discontinuous (Figure 1b).

The magnet pictured in Figure 1a can be considered a "device" that acts on an incident beam. The effect of this device on a two-state spin system is representable by a two-by-two matrix similar to the matrices introduced in Chapter 5 to represent devices that act on photon polarization. The elements of the matrix that represents this particular device may be determined in the usual way: beams of atoms in various spin states are passed through the device and the state of the atoms in the output beam is tested using appropriate projectors.

To predict the values of the matrix elements, we can make use of a semi-classical argument. The behavior of a magnetic dipole in a uniform field was analyzed from a classical viewpoint in Box 1 of Chapter 7. It was concluded there that the angular momentum and magnetic moment vectors precess about the field direction at an angular frequency given by Eq. 1.



a. Schematic diagram of experiment.



b. Idealized field plot as a function of position along the beam.

Figure 1. Idealized experiment to examine the precession of a two-state spin system in a uniform magnetic field.

$$\omega_p = \mu B/J = geB/(2mc) = g\omega_L \tag{1}$$

where J is the magnitude of the angular momentum* and

$$\omega_L = eB/(2mc) \tag{2}$$

is called the Larmor frequency. The factor g was introduced in order to allow for possible departures from the simplest classical model (g = 1) in which both angular momentum and magnetic moment arise from a point charge circulating in an orbit. Let Δt denote the time that any given atom requires to pass through the field of Figure 1. Then the total angle of precession $\Delta\phi$ of the angular momentum vector in this time is, according to the classical picture,

$$\Delta\phi = g\omega_L \Delta t \tag{3}$$

There is no reason to believe that classical precession correctly describes the behavior of a two-state spin system in a uniform magnetic field. In Chapter 7 we found classical predictions for this system to be undependable and often exactly wrong. Nevertheless, both in Chapter 7 and here, the classical picture is useful in setting up a quantum prediction by analogy that can then be tested for correctness on its own merits. In the present case, the classical precession model is readily translated into the language of quantum states. The analysis suggests that if the atoms in the incident beam in the experiment of Figure 1 are in the state

* In Chapter 7, L was used for the orbital angular momentum of the electron. Total angular momentum of an atom may include also spin angular momentum of electrons. The symbol $J = j\hbar$ is the conventional symbol for the total angular momentum from all sources.

$|\theta_o, \phi_o\rangle$, then the atoms in the output beam will be in the state

$$|\theta_f, \phi_f\rangle = |\theta_o, \phi_o + \Delta\phi\rangle \tag{4}$$

where $\Delta\phi$ is given by Eq. 3. (See Figure 2.) This result is in fact quantum mechanically correct. We remark in passing that it has already been employed in Chapter 7, in connection with the production of states of longitudinal polarization by passing a beam through a uniform magnetic field.

Experimental verification of (4) is sometimes direct and sometimes indirect. Observe that for the special case in which the initial angular momentum is, according to the classical model, lined up either along or against the field direction ($\theta_o = \pi$ or 0), there is no precession. The corresponding quantum statement, verified directly by experiment, is that when the incident atoms are in the quantum state $|+z\rangle$ or $|-z\rangle$, the atoms in the output beam are in the same state as those in the input. If A represents the device, this experiment shows that

$$\langle -z|A|+z\rangle = \langle +z|A|-z\rangle = 0 \tag{5}$$

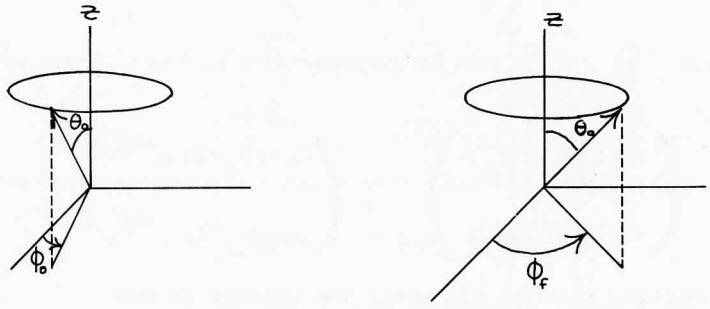


Figure 2. Classical picture of precession of a magnetic dipole in a z-oriented magnetic field. Straight arrow represents the angular momentum of the dipole.

Consider another special case: the input atoms are in the state $|+x\rangle$ and values of B and Δt (Eq. 3) are chosen so as to make the precession angle $\Delta\phi = \pi$. Then the atoms in the output beam are predicted to be in state $|-x\rangle$. For such a device

$$\langle +x|A|+x\rangle = \langle -x|A|-x\rangle = 0 \quad \text{(special choice of B and } \Delta t) \tag{6}$$

Many practical difficulties lie in the way of verifying directly predictions like (6). For one thing the changes in field direction between $\pm x$ -projectors and device A must be "sufficiently abrupt" as discussed in Section 4 of Chapter 7. In addition, the precession frequency for atoms is, under normal laboratory conditions, quite high. For a field of one gauss, for example, Eq. 1 with m equal to the mass of the electron gives $\omega_p \approx 9 \times 10^6$ radians per second (assuming g = 1). Typical velocities for atomic beams are of the order of 10^5 centimeters per second. Therefore in traversing a field region only 5 centimeters long an atomic moment precesses more than 70 times. Under such conditions even a small spread in the velocity of the beam particles or a slight nonuniformity in the field will result in an output beam in

which the particles are in different polarization states. To make the precession of the order of one revolution requires a combination of extremely small uniform field and very high and uniform beam velocity--a combination difficult to attain in practice. Controlled precession of the kind discussed here is achieved with neutrons from nuclear reactions. Success in this case depends upon the small magnetic moment of the neutrons (2000 times smaller than that of atoms) and their very high and well-controlled velocity (energies in the MeV range). For atoms, firm evidence to confirm the predictions (4) is provided by "resonance" experiments to be discussed in a later chapter. Such experiments verify, indirectly but conclusively, the picture of "precessing states" presented in the present chapter.

It is not hard to find the matrix A that represents the device in Figure 1. Equation 5 indicates that in the z-representation, A must be diagonal. Beams in state $|+z\rangle$ and in state $|-z\rangle$ come through the device undiminished in intensity, so the diagonal elements have magnitude unity. We can verify directly that the desired matrix is

$$A = \begin{pmatrix} e^{-i\Delta\phi/2} & 0 \\ 0 & e^{+i\Delta\phi/2} \end{pmatrix} \qquad \text{z-representation} \qquad (7)$$

The beam in initial state $|\theta_o, \phi_o\rangle$ can be represented by the column vector (Eq. 21 of Chapter 7)

$$\begin{pmatrix} \langle +z|\theta_o, \phi_o\rangle \\ \langle -z|\theta_o, \phi_o\rangle \end{pmatrix} = \begin{pmatrix} \cos(\theta_o/2)\, e^{-i\phi_o/2} \\ \sin(\theta_o/2)\, e^{+i\phi_o/2} \end{pmatrix} \qquad (8)$$

Acting on this column vector, matrix (7) gives the column vector

$$\begin{pmatrix} \cos(\theta_o/2)\, e^{-i(\phi_o + \Delta\phi)/2} \\ \sin(\theta_o/2)\, e^{+i(\phi_o + \Delta\phi)/2} \end{pmatrix} \qquad (9)$$

which we recognize as representing the state $|\theta_o, \phi_o + \Delta\phi\rangle$. This proves that (7) is indeed the correct matrix.

3. The time-dependent state vector.

The discussion has thus far involved only a comparison of the final with the initial state. As an alternative, we may try to follow the behavior of the particles as a function of time during the period in which they are subject to the magnetic field. In the classical version of the problem this is accomplished merely by writing the equation that describes the precession as a function of time. If a specific atom enters the field region at $t = 0$ with its angular momentum pointing in the direction θ_o, ϕ_o, then at time t the angular momentum points in the direction $\theta(t), \phi(t)$ given by the equations

$$\begin{aligned} \theta(t) &= \theta_o \\ \phi(t) &= \phi_o + \omega_p t \end{aligned} \qquad (10)$$

where ω_p is given by Eq. 1.

The quantum mechanical equivalent of Eqs. 10 is a time-dependent state vector that gradually evolves from the initial state to a later state. Let $|\psi(t)\rangle$ denote such a time-dependent state vector. In the present problem the time dependence appears explicitly only in the angle ϕ. We can write, therefore,

$$|\psi(t)\rangle = |\theta_o, \phi(t)\rangle = |\theta_o, \phi_o + \omega_p t\rangle \tag{11}$$

$$= |+z\rangle \cos\frac{\theta_o}{2}\, e^{-i(\phi_o + \omega_p t)/2} + |-z\rangle \sin\frac{\theta_o}{2} e^{+i(\phi_o + \omega_p t)/2}$$

The latter expression is in the z-representation. The state vector (11) represents the state of an atom while it is passing through the magnetic field, in the sense that if a measurement is carried out at any time t, (11) can be used to predict the results. Notice that the time is defined individually for each atom: in the present example, $t = 0$ denotes the instant at which the atom enters the field region. Equation 11 shows explicitly that the amplitudes of a time-dependent state vector with respect to any set of basis states are themselves time dependent. For the special case of an x-polarized incident beam ($\theta_o = \pi/2$, $\phi_o = 0$), the amplitudes of state (11) in the z representation are

$$\begin{aligned} \langle +z|\psi(t)\rangle &= \frac{1}{\sqrt{2}}\, e^{-i\omega_p t/2} \\ \langle -z|\psi(t)\rangle &= \frac{1}{\sqrt{2}}\, e^{i\omega_p t/2} \end{aligned} \tag{12}$$

Alternatively, in the x-representation the same state has the amplitudes

$$\begin{aligned} \langle +x|\psi(t)\rangle &= \cos(\omega_p t/2) \\ \langle -x|\psi(t)\rangle &= i \sin(\omega_p t/2) \end{aligned} \tag{13}$$

The absolute squares of the amplitudes in (12) and (13) specify the relative intensities observed if the beam is passed through the corresponding analyzer after spending time t in the field region. Notice that the quantities $|\langle +x|\psi(t)\rangle|^2$ and $|\langle -x|\psi(t)\rangle|^2$ are themselves time varying, whereas the time dependence of $\langle +z|\psi(t)\rangle$ and $\langle -z|\psi(t)\rangle$ consists only of phase factors whose absolute magnitude is unity for any time whatsoever. This reflects the fact that the results of one possible experiment on the beam (analyzing it for $\pm x$ polarization) are time-dependent, whereas another possible experiment (analysis for $\pm z$ polarization) gives results that are time-independent.

4. Stationary states

A role of particular importance in quantum physics is played by those states which, for a given physical system subject to specified forces, lead to time-independent results for all possible measurements. Such states are called stationary states. It is not obvious that any such states ought to exist at all. The general conditions under which a system can take up stationary states will be discussed in a later chapter. Stationary states can exist whenever external in-

fluences on the system are time-independent. In the present instance the external influence is the torque that acts, independent of time, on a spin in a uniform magnetic field.

For a two-state spin system in a uniform magnetic field, the stationary states are those two states polarized respectively along the field direction and opposite to the field direction. With the field in the z direction, the complete time-dependent state vectors for the stationary states can be written as follows

$$\begin{aligned} |\psi_+(t)\rangle &= |+z\rangle e^{-i\omega t} \\ |\psi_-(t)\rangle &= |-z\rangle e^{i\omega t} \end{aligned} \tag{14}$$

where

$$\omega = \omega_p/2 = \frac{1}{2}\, g\omega_L \tag{15}$$

Results of all experiments on either of these states are independent of the time the atoms spend in the magnetic field. In contrast, if atoms are in a state that represents polarization along an axis in any direction other than that of the field, then the result of some experiments are time-dependent--i.e., they depend upon the time the atoms spend in the uniform magnetic field.

Clearly a state is stationary according to the definition given above if its time dependence is contained in a single exponential of the form $e^{\pm i\omega t}$. Indeed, we shall see later that an imaginary exponential linear in the time is also the most general form for the time dependence of any stationary state. Even now we can see two arguments that support this important result:

1. Any amplitude whose magnitude is independent of the time can contain the time only in the imaginary argument of an exponential: $e^{if(t)}$, where the function $f(t)$ is real.
2. The change in the phase factor between time t_1 and later time t_2

$$e^{i[f(t_2) - f(t_1)]}$$

should depend only on the elapsed time $t_2 - t_1$. This implies that the function $f(t)$ is a linear function of time.

The coefficient of t in the exponentials of Eq. 14 is proportional to the potential energy of the spins in the magnetic field. The classical expression for energy of orientation of a dipole in a magnetic field is

$$E = -\vec{\mu} \cdot \vec{B} \tag{16}$$

For the special cases in which the dipole is aligned along or against the direction of the magnetic field this energy takes on the values

$$E_+ = +\mu B \tag{17a}$$

for $\vec{\mu}$ antiparallel to $\vec{B}$ (state $|+z\rangle$), and the value

$$E_- = -\mu B \tag{17b}$$

for $\vec{\mu}$ parallel to $\vec{B}$ (state $|-z\rangle$). Using equations (1) and (15) we can write (17a, b) as

$$E_\pm = \pm J\omega_p = \pm 2J\omega = \pm 2j\hbar\omega \tag{18}$$

where $j = J/\hbar$ is the angular momentum in units of $\hbar = h/2\pi$.

In Chapter 7 we hinted at the fact that a two-state spin system represents angular momentum $J = \hbar/2$ or $j = J/\hbar = 1/2$. Two-state spin systems are called spin one-half systems. With $j = 1/2$, equation 18 becomes

$$E = \hbar\omega \tag{19}$$

This equation is the same as the one that relates photon energy to the classical frequency of an electromagnetic wave. The equation has important repercussions throughout quantum physics. Whenever a stationary state is found for a physical system, its dependence upon time is always related to its energy through the phase factor

$$(\text{phase factor for stationary state}) = e^{-i(E/\hbar)t} \tag{20}$$

Conversely, whenever a given quantum state is associated with a definite energy, it is a stationary state and its time dependence is of the exponential form (20). The parameter $\omega = E/\hbar$ sometimes--but not always--corresponds to an analogous classical frequency. In quantum physics this frequency appears in the phase of an amplitude and can be detected in practice only when amplitudes for different stationary states interfere.

In addition to the potential energy given by Eq. 16, the atomic beam particles also possess a kinetic energy $p^2/2m$. Strictly, speaking, this energy should also be included in the time dependence of the states. However, since the kinetic energy is (very nearly) the same for all possible spin states of the beam,* its effect is simply to multiply any state vector by the overall phase factor $e^{-i(p^2/2m)t/\hbar}$. This overall phase factor does not affect the result of any measurement and may therefore be omitted. In Chapter 10 we will consider a superposition of two beams with different momenta; in that case the kinetic energy contribution to the time-dependent phase cannot be omitted and indeed leads to a description of moving pulses of particles.

A similar remark applies to the photon polarization states that we have studied earlier: A time dependence of the form (19) ought to be included in all the state vectors and amplitudes that have been written. But this time dependence is just $e^{-i\omega t}$, where ω is the classical frequency of the photons. We have implicitly assumed that the frequency is well defined and has the same value for all the photons in a given experiment. Therefore the same time-dependent factor appears in all amplitudes, and disappears when the absolute square is finally taken to evaluate a probability. For experiments involving photons of the same energy, it is unnecessary to carry the time-dependent factor along. When amplitudes corresponding to different energies interfere, then the time dependence is important.

The universal applicability of Eq. 19 to stationary states supports the postulate that j has the value $1/2$ for a two-state spin system. A direct macroscopic measurement of the angular momentum of atoms, similar to the one discussed in Chapter 2 for photons, is in

---

* That this is in fact true may be verified by measuring the speed of the particles in the beams separated by an analyzer.

principle possible, but is a very difficult experiment to carry out. Experiments with atomic spectra provide convincing proof that $j = 1/2$ for such systems.

The stationary states of a two-state spin system in a uniform magnetic field can be written, using (14) and (20), in the form

$$|\psi_{\pm}(t)\rangle = |\pm z\rangle e^{-i(E_{\pm}/\hbar)t} \tag{21}$$

Look again at the difference between a stationary state and one that is not stationary. Consider a beam of initially +x-polarized atoms entering a region of uniform magnetic field (Figure 1). The classical prediction is that each magnetic dipole will remain perpendicular to the magnetic field while precessing in the xy plane. Quantum predictions are summarized in the state vector. Equations 13 and 15 allow us to write the state vector $|\mathcal{E}(t)\rangle$ for the system in the x-representation as

$$|\mathcal{E}(t)\rangle = |+x\rangle \cos \omega t + |-x\rangle \, i \sin \omega t \tag{22}$$

Clearly this is not a stationary state. In particular, the projection probability into the state +x depends on the time t that the spins have spent in the uniform magnetic field.

There is an important difference between the classical and quantum points of view concerning the energy associated with a state such as (22). Classically, a dipole precessing in the xy plane has zero energy, since it always points in a direction perpendicular to the z-oriented field ($E = -\vec{\mu} \cdot \vec{B} = 0$ for this case). Quantum mechanically we have to consider how the energy is to be measured. At the present stage of the development, the only means available to measure the potential energy of an atom in a magnetic field is to use a z-analyzer with axis along the field direction. The atom then is measured to have energy $E_+ = +\mu B$ or $E_- = -\mu B$ depending on whether it is found, respectively, in the upper or lower beam when it leaves the analyzer. As defined by this procedure, the result of <u>every</u> potential energy measurement is either E_+ or E_-. For the particular case of the state (22), we can predict the result of such a measurement by expressing the state in terms of the stationary states $\pm z$:

$$|\mathcal{E}(t)\rangle = |+z\rangle \frac{1}{\sqrt{2}} e^{-i(E_+/\hbar)t} + |-z\rangle \frac{1}{\sqrt{2}} e^{-i(E_-/\hbar)t} \tag{23}$$

Equation 23 shows that the probabilities of the two possible results of an energy measurement are each 1/2, independent of the time of the measurement. Therefore the <u>average</u> value of the energy (the "expectation value") is zero for state $|\mathcal{E}(t)\rangle$, in accord with the classical interpretation. This average is zero even though individual measurements never yield the result zero: atoms come either through channel +z (energy E_+) or through channel -z (energy E_-). This is in marked contrast to the results of energy measurement on a beam in a stationary state. For example, a beam initially in state +z after passing through a uniform magnetic field (state vector $|\psi_+(t)\rangle = |+z\rangle e^{-i(E_+/\hbar)t}$) will all be transmitted in the upper channel of a z-analyzer: an energy measurement on this state <u>always</u> gives the energy E_+.

In this chapter we have returned--for the first time since Chapter 1--to energy as a parameter of a quantum state. Transitions between quantum states of different energy are often

accompanied by the emission or absorption of photons with discrete energies. The same is true for the two-state spin system in a magnetic field. A photon of angular frequency $(E_+ - E_-)/\hbar = \omega_p$ can be absorbed or emitted, as will be discussed more fully in a later chapter. For laboratory magnetic fields (a few thousand gauss) this photon is in the range of energies of radio waves. The fact that only one such frequency is detected is further evidence that there are only two possible states of definite energy--two stationary states.

In summary, the characteristic of a system in a stationary state is that the result of every measurement on it (including an energy measurement) always yields the same result independent of the time at which the measurement takes place.