

Chapter 1

Basic Concepts

1.1 Spin States and Density Matrix of Spin-1/2 Particles

1.1.1 Pure Spin States

In order to become familiar with the basic concepts of density matrix theory we will begin by considering the problem of describing the spin states of spin-1/2 particles. First of all we will review some results of the quantum mechanical theory of experiments with Stern–Gerlach magnets and then in the following sections we will reinterpret the results and discuss them in more detail.

Consider a beam of spin-1/2 particles (for example, hydrogen atoms) which passes through a Stern–Gerlach magnet which has its field gradient aligned along the z direction with respect to a fixed coordinate system x, y, z (Fig. 1.1). In general the beam will split vertically into two parts each of which corresponds to one of the two possible eigenvalues of the component S_z of the spin operator \mathbf{S} ($m = \pm 1/2$). If one of the beams is eliminated, for example, the lower one as in Fig. 1.1, then the emerging particles will be in a state which corresponds to only one of the eigenvalues; in the case of the apparatus in Fig. 1.1 this would be $m = +1/2$. Similarly, if the apparatus is rotated in such a way that its field gradient points in the direction z' , the emerging particles will be in a state which is described by the quantum number $m' = +1/2$, where m' is the eigenvalue of the operator $S_{z'}$, the component of \mathbf{S} in the z' direction.

If the incident beam is such that it contains particles which are in a state with $m = +1/2$ only, then the beam will pass through the apparatus shown in Fig. 1.1 without any loss of intensity. In all other cases part of the beam will be blocked off and the emerging beam will be less intense than the incident one. However, by tilting the apparatus at various angles about z it may be possible to find an orientation of the magnet which allows the whole beam to be transmitted. For example, if an incident beam contains only a spin component corresponding to $m' = +1/2$ in the frame z' , it would be attenuated by the Stern–Gerlach apparatus in Fig. 1.1. If the

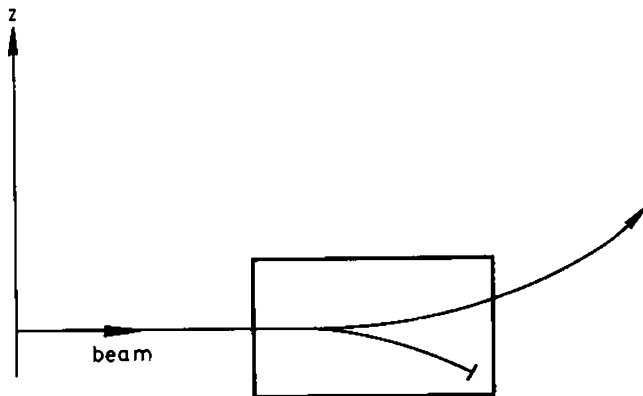


Fig. 1.1 Stern–Gerlach filter

magnet were rotated so that its field gradient lay along z' , then the beam would be completely transmitted. In this case all particles are deflected in the same way; they behave identically in this particular experiment. This enables the following (preliminary) definition to be made:

- If it is possible to find an orientation of the Stern–Gerlach apparatus for which a given beam is *completely* transmitted, then we will say that the beam is in a *pure spin state*.

In terms of the semiclassical vector model a beam of particles with definite quantum number $m = +1/2$ can be described by considering the spin vector of each particle to precess around the direction of the z axis such that its projection on the z axis has the value of $+1/2$ (Fig. 1.2a). In this case the particles are said to have “spin up.” A similar interpretation holds for the case of $m = -1/2$ (Fig. 1.2b) and the spins of particles in eigenstates of the operator $S_{z'}$ will, by analogy, precess around the z' direction. In the case of a pure spin state the spin vectors of the particles precess around a unique direction which is parallel to the direction of alignment of the Stern–Gerlach apparatus when it allows the beam to pass through unattenuated.

If the state of a given beam is known to be pure then the joint state of all particles can be represented in terms of one and the same state vector $|\chi\rangle$. This is an important point and we will illustrate it with some examples. If a beam of particles passes completely through a Stern–Gerlach apparatus oriented in the z direction then we will say that *all* particles in the beam are in identical spin states with quantum number $m = 1/2$ with respect to z , or that all particles have spin up with respect to z . We describe this state by assigning the state vector $|\chi\rangle = | + 1/2 \rangle$ to the *whole* beam. Similarly, a beam of particles with $m = -1/2$ will be characterized by $|\chi\rangle = | - 1/2 \rangle$. In the usual Pauli representation the state vectors are represented by two-dimensional column vectors,

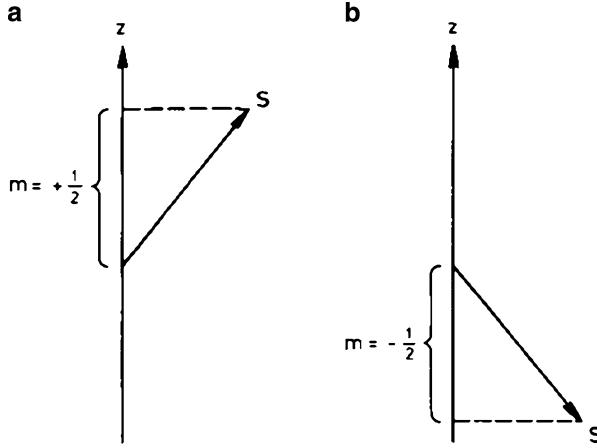


Fig. 1.2 (a) Spin “in the z direction”; (b) spin “in the $-z$ direction”

$$\left| +\frac{1}{2} \right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad \left| -\frac{1}{2} \right\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (1.1a)$$

and the adjoint states by the row vectors

$$\left\langle +\frac{1}{2} \right| = (1, 0); \quad \left\langle -\frac{1}{2} \right| = (0, 1) \quad (1.1b)$$

In general, for a beam that emerges from a Stern–Gerlach apparatus which has its magnet pointing in the z' direction, all particles in the beam are in a state with definite spin quantum number $m' = 1/2$ defined with respect to z' as quantization axis. The joint state of all particles will be described by the state vector $|\chi\rangle = | +1/2, z'\rangle$.

A general spin state $|\chi\rangle$ can always be written as a linear superposition of two basis states, for example, the states $|\pm 1/2\rangle$:

$$|\chi\rangle = a_1 \left| +\frac{1}{2} \right\rangle + a_2 \left| -\frac{1}{2} \right\rangle \quad (1.2)$$

In the representation (1.1) this is equivalent to

$$|\chi\rangle = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad (1.3a)$$

The adjoint state is represented by the row vector

$$\langle \chi| = (a_1^*, a_2^*) \quad (1.3b)$$

where the asterisk denotes the complex conjugate.

The state $|\chi\rangle$ is normalized such that

$$\langle\chi|\chi\rangle = |a_1|^2 + |a_2|^2 = 1 \quad (1.4)$$

A pure spin state can be characterized either by specifying the direction in which the spins are pointing (for example, by giving the polar angles of this direction in our fixed coordinate system) or, alternatively, by specifying the coefficients a_1 and a_2 in the expansion (1.2). In the following section we will discuss the relation between these two descriptions and derive an explicit representation for the coefficients a_1 and a_2 .

An apparatus of the type shown by Fig. 1.1 acts as a *filter*, because irrespective of the state of the beam sent through it, the emerging beam is in a definite spin state which is defined by the orientation of the magnet. Passing a beam through the filter can therefore be regarded as a method of preparing a beam of particles in a pure state.

1.1.2 The Polarization Vector

In order to discuss the description of pure spin states in greater detail we will now introduce a vector \mathbf{P} , called the *polarization vector*, which has components defined as expectation values of the corresponding Pauli matrices:

$$P_i = \langle\sigma_i\rangle \quad (1.5)$$

($i = x, y, z$). In the case of a pure state these expectation values are defined by the relations.

$$\langle\sigma_i\rangle = \langle\chi|\sigma_i|\chi\rangle \quad (1.6)$$

In the representation (1.1) the Pauli matrices have the form

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (1.7)$$

The expectation values (1.6) may then be calculated by applying (1.3a), (1.3b), and (1.7), treating the row and column vectors as one-dimensional matrices and applying the rules of matrix multiplication. In order to see the significance of the polarization vector we will now consider a few examples.

A beam of particles in the pure state $|+1/2\rangle$ has a polarization vector with components

$$P_x = (1, 0) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0$$

$$P_y = (1, 0) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0 \quad (1.8a)$$

$$P_z = (1, 0) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1$$

Similarly we find for an ensemble of particles in the pure state $| - 1/2 \rangle$ the polarization vector has components

$$P_x = 0, \quad P_y = 0, \quad P_z = -1 \quad (1.8b)$$

Thus the states $| + 1/2 \rangle$ and $| - 1/2 \rangle$ are characterized by polarization vectors of unit magnitude pointing in the $+z$ and $-z$ directions, respectively. The states $| + 1/2 \rangle$ and $| - 1/2 \rangle$ can therefore be said to be states of opposite polarization.

Consider now the general pure state (1.2). It will be convenient to give first of all a parametrization of the coefficients a_1 and a_2 . These are complex numbers corresponding to four real parameters specifying the magnitudes and phases. The overall phase of the state (1.2) has no physical significance and can be chosen arbitrarily, for example by requiring a_1 to be real. From this condition and the normalization (1.4) it follows that the general pure spin state (1.2) is completely specified by two real numbers. It will therefore be convenient to introduce two parameters θ and δ defined by

$$a_1 = \cos \frac{\theta}{2}, \quad a_2 = e^{i\delta} \sin \frac{\theta}{2} \quad (1.9)$$

where δ is the relative phase of the coefficients. Using (1.9), (1.3a) becomes

$$|\chi\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ e^{i\delta} \sin \frac{\theta}{2} \end{pmatrix} \quad (1.10)$$

In order to see the physical significance of the parameters θ and δ consider the polarization vector associated with the state (1.10). We obtain

$$\begin{aligned} P_x &= \left(\cos \frac{\theta}{2}, e^{-i\delta} \sin \frac{\theta}{2} \right) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \cos \frac{\theta}{2} \\ e^{i\delta} \sin \frac{\theta}{2} \end{pmatrix} \\ &= \sin \theta \cos \delta \end{aligned} \quad (1.11a)$$

$$P_y = \sin \theta \sin \delta \quad (1.11b)$$

$$P_z = \cos \theta \quad (1.11c)$$

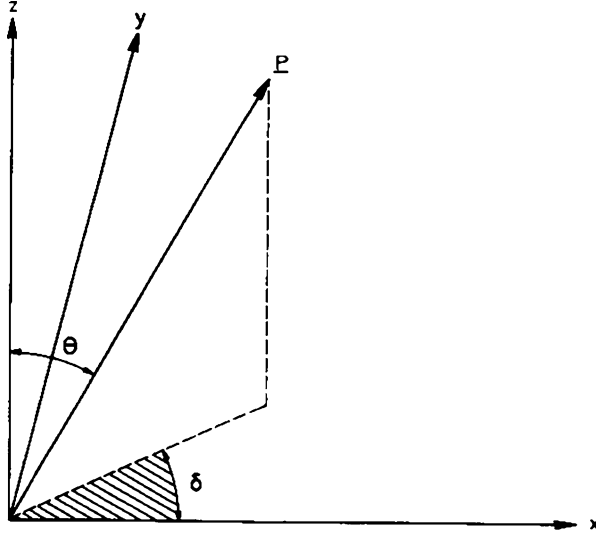


Fig. 1.3 Direction of \mathbf{P}

The polarization vector (1.11) has unit magnitude

$$|\mathbf{P}| = (P_x^2 + P_y^2 + P_z^2)^{1/2} = 1 \quad (1.12)$$

From (1.11) it follows that the parameters θ and δ can be interpreted as the polar angles of \mathbf{P} : θ is the angle between \mathbf{P} and the z axis and the relative phase δ specifies the azimuth angle of \mathbf{P} (Fig. 1.3).

A second coordinate system x' , y' , z' can be chosen in such a way that the z' axis is parallel to \mathbf{P} . Taking z' as quantization axis we have in the primed system $P_{x'} = 0$, $P_{y'} = 0$, $P_{z'} = 1$; that is, all particles have spin up with respect to z' . The direction of the polarization vector is therefore the direction “in which the spins are pointing.” If we send the beam through a Stern–Gerlach filter oriented parallel to \mathbf{P} the whole beam will pass through.

Equations 1.10 and 1.11 enable explicit spin functions to be constructed for any pure state. For example, a given beam of particles may be in a pure state with spins pointing in the x direction of our fixed coordinate system. In this case the corresponding polarization vector points in the x direction and, consequently, has polar angles $\theta = 90^\circ$, $\delta = 0$. From (1.10) the state vector can be seen to be

$$\left| +\frac{1}{2}, x \right\rangle = \frac{1}{2^{1/2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (1.13a)$$

A beam of particles with “spin down” with respect to the x axis has a polarization vector pointing in the $-x$ direction and is characterized by the angles $\theta = 90^\circ$,

$\delta = 180^\circ$. The state vector is represented by

$$\left| -\frac{1}{2}, x \right\rangle = \frac{1}{2^{1/2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (1.13b)$$

Similarly, the state vector of particles with “spin up” (“spin down”) with respect to the y axis is represented by the column vectors

$$\left| +\frac{1}{2}, y \right\rangle = \frac{1}{2^{1/2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \quad (1.13c)$$

and

$$\left| -\frac{1}{2}, y \right\rangle = \frac{1}{2^{1/2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \quad (1.13d)$$

It should be noted that the four states (1.13) are constructed by superposing the states $| + 1/2 \rangle$ and $| - 1/2 \rangle$ with the same magnitude $|a_1| = |a_2| = 1/2^{1/2}$ but with different relative phases. The corresponding polarization vectors have the same angle θ but different azimuth angles depending on the relative phase existing between the states $| \pm 1/2 \rangle$.

1.1.3 Mixed Spin States

Pure spin states are not the most general spin states in which an ensemble of particles can be found. Suppose, for example, that two beams of particles have been prepared *independently*, one in the pure state $| + 1/2 \rangle$ and the other one in the pure state $| - 1/2 \rangle$. By “independently” we mean that no definite phase relation exists between the two beams (this point will be clarified later). The first beam may consist of N_1 particles, the second one of N_2 particles. If the polarization state of the combined beam is investigated by sending it through a Stern–Gerlach filter in various orientations it will be found that it is not possible to find such an orientation of the filter which allows the whole beam to be transmitted. It follows that, by definition, the joint beam is not in a pure spin state.

- States which are not pure are called *mixed states* or *mixtures*.

We now have to consider the problem of describing the state of the joint beam. Clearly, it is not possible to characterize the state of the beam in terms of a single state vector $|\chi\rangle$ since associated with any of these states there is necessarily a direction in which *all* spins are pointing: the direction of the polarization vector. If the Stern–Gerlach filter were placed in this orientation the whole beam would have to be transmitted. Since no such orientation exists it is not possible to describe a mixture by a single state vector.

In particular, the mixture cannot be represented by a linear superposition of the states $|+1/2\rangle$ and $|-1/2\rangle$ representing the two constituent beams. In order to construct such a linear superposition it is necessary to know the magnitudes and relative phase δ of the relevant coefficients a_1 and a_2 . The absolute squares $|a_1|^2$ and $|a_2|^2$ are probabilities W_1 and W_2 of finding a particle in the state $|+1/2\rangle$ and $|-1/2\rangle$, respectively. In the case of the mixture under discussion these probabilities are known ($W_1 = N_1/N$ and $W_2 = N_2/N$ with $N = N_1 + N_2$) and can be used to determine the magnitudes of the coefficients ($W_1 = |a_1|^2$, $W_2 = |a_2|^2$). The important point is that the constituent beams have been prepared independently. So there is no definite phase relation between the two beams, and without a definite phase δ it is not possible to construct a well-defined state vector $|\chi\rangle$ describing the joint beam.

A mixture has to be described by specifying the way in which it has been prepared. For example, the joint beam under discussion is characterized by saying that N_1 particles have been prepared in the state $|+1/2\rangle$ and N_2 in the state $|-1/2\rangle$ independently of each other. This statement contains all the information we have obtained about the mixture.

Let us continue discussion of our example by calculating the polarization vector associated with the total beam. \mathbf{P} is obtained by taking the statistical average over the separate beams:

$$P_i = W_1 \left\langle \frac{1}{2} |\sigma_i| \frac{1}{2} \right\rangle + W_2 \left\langle -\frac{1}{2} |\sigma_i| - \frac{1}{2} \right\rangle$$

which gives

$$P_x = 0, \quad P_y = 0, \quad P_z = W_1 - W_2 = \frac{N_1 - N_2}{N} \quad (1.14)$$

It should be noted that the polarization vector (1.14) has a magnitude which is less than 1 and is proportional to the difference of the population numbers of the two states, $|+1/2\rangle$ and $|-1/2\rangle$.

More generally, if a beam is prepared by mixing N_a particles in the state $|\chi_a\rangle$ and N_b in the state $|\chi_b\rangle$ then the components of the polarization vector are given by the statistical average over the independently prepared beams:

$$P_i = W_a \langle \chi_a | \sigma_i | \chi_a \rangle + W_b \langle \chi_b | \sigma_i | \chi_b \rangle \quad (1.15)$$

$$= W_a P_i^{(a)} + W_b P_i^{(b)} \quad (1.16a)$$

with $W_a = N_a/N$, $W_b = N_b/N$, and where $P_i^{(a)}$ and $P_i^{(b)}$ are the polarization vectors associated with the constituent beams [see, (1.6)]. Equation 1.15 can be rewritten in vector notation as

$$\mathbf{P} = W_a \mathbf{P}^{(a)} + W_b \mathbf{P}^{(b)} \quad (1.16b)$$

Since $|\mathbf{P}^{(a)}| = 1$, $|\mathbf{P}^{(b)}| = 1$ the magnitude of P is determined by

$$\begin{aligned}
 P^2 &= (W_a \mathbf{P}^{(a)} + W_b \mathbf{P}^{(b)})^2 \\
 &= W_a^2 (P^{(a)})^2 + W_b^2 (P^{(b)})^2 + 2W_a W_b \mathbf{P}^{(a)} \cdot \mathbf{P}^{(b)} \\
 &\leq W_a^2 + W_b^2 + 2W_a W_b \\
 &= (W_a + W_b)^2 = 1
 \end{aligned} \tag{1.17}$$

since the scalar product $\mathbf{P}^{(a)} \cdot \mathbf{P}^{(b)}$ of two different unit vectors is less than 1.

The equality sign in (1.17) applies if $\mathbf{P}^{(a)} \cdot \mathbf{P}^{(b)} = 1$, that is, if the two beams have identical polarization vectors. In this case both constituent beams are in the same spin state described by (1.10) and (1.11) and the joint beam is in a pure state. Vice versa, if two beams are mixed in identical spin states the resulting beam consists of particles in identical spin states and is therefore characterized by a polarization vector of unit magnitude. The above arguments may be easily generalized to cases of mixtures which consist of more than two beams.

We therefore have the following result: The magnitude of the polarization vector is bounded such that

$$0 \leq |\mathbf{P}| \leq 1 \tag{1.18}$$

The maximum possible value $|\mathbf{P}| = 1$ is obtained if (and only if) the beam under consideration is in a pure state, whereas mixtures necessarily have a polarization vector of less than unit magnitude.

This result once more illustrates the basic property of a pure spin state: all the particles are in identical states with *all* the spins pointing in the same direction, the direction of \mathbf{P} .

Henceforward we will refer to states with $|\mathbf{P}| > 0$ as *polarized* and to beams with $|\mathbf{P}| = 0$ as *unpolarized*. Pure states ($|\mathbf{P}| = 1$) will be called *completely polarized*.

1.1.4 Pure Versus Mixed States

Before proceeding further with any analysis it is important to have a clear understanding of the concepts of pure and mixed states. We will therefore consider both types of states again from a different viewpoint. Consider the following problem. A beam of particles which is completely polarized in the y direction and hence can be represented by the state vector (1.13c)

$$\left| +\frac{1}{2}, y \right\rangle = \frac{1}{2^{1/2}} \left(\left| +\frac{1}{2} \right\rangle + i \left| -\frac{1}{2} \right\rangle \right) \tag{1.19a}$$

is sent through a Stern–Gerlach filter oriented in the z direction. What will happen? It is a familiar result of quantum mechanics that, although we know that any particle

in the beam is in the state $| + 1/2, y \rangle$, it is impossible to predict whether a given single particle will pass through the filter. This is because a system which is to be measured is in general disturbed by the act of measurement. In this case the measuring apparatus (the filter) changes in a completely uncontrollable way the state of the incident particles; that is, it is only possible to predict the probability that a particle will be admitted by the filter (and emerge in the state $| + 1/2 \rangle$) or be blocked off. From (1.19a) it can be seen that the probability is $1/2$ for each case. The only case in which it is possible to predict with complete certainty whether a given particle will pass through a filter is that where the filter is oriented in the y direction all particles will pass unhindered. In general, however, the measuring process can only be described through the use of statistics.

Because of this a linear superposition state, such as (1.19a), must be interpreted as follows. Before any measurement all the particles are in identical states represented by the vector (1.19a), and all particles have the same quantum number $m' = 1/2$ defined with respect to y as the quantization axis. The quantum number m , defined with respect to the z axis, is completely undefined in the superposition (1.19a) in the sense that *any* particle in the beam has an equal probability of passing through a z -oriented filter or being blocked off. [Roughly speaking, it can be said that the particles in the superposition state (1.19a) do not “know” their m value.] If the beam is sent through a filter oriented parallel to the z axis the interaction with the apparatus changes the state of the beam and forces the particles into one of the eigenstates.

Consider now a mixture of

$$\begin{aligned} N_1 &= N/2 \text{ particles in the state } | + 1/2 \rangle \\ N_2 &= N/2 \text{ particles in the state } | - 1/2 \rangle \end{aligned} \quad (1.19b)$$

with both subbeams prepared independently. From (1.14) it can be seen that the resulting beam is unpolarized. If this beam is passed through a Stern–Gerlach filter oriented along the z axis the transmitted beam will have half of the incident intensity. In this particular experiment the mixture (1.19b) and the pure state (1.19a) give the same result; however, it is for a different reason. Whereas in the case of the state $|1/2, y\rangle$ all particles in the beam are in one and the same state, there is less information about the mixture (1.19b) since it is only known that any particle has an equal probability of being in the state $| + 1/2 \rangle$ or $| - 1/2 \rangle$. In this sense the state of the mixture is *incompletely determined*. When passing through the filter the particles with $m = -1/2$ will be blocked off and hence only that half of the beam corresponding to the $| + 1/2 \rangle$ -component beam will be transmitted.

The above example illustrates that statistics must be used in order to describe the *initial* state of the mixture; the state of the particles is not known with certainty; that is, we cannot assign a single-state vector to the mixed beam.

In conclusion, it can be seen that in the description of spin- $1/2$ particles statistics enters in two ways. First of all, statistical methods must be used because of the uncontrollable perturbation of states by any measuring apparatus. Secondly, when

dealing with mixtures, it is only known that the particles can be in any one of several spin states. A statistical description must be applied because of the lack of information available on the system. It was primarily for the purpose of describing this latter case that the density matrix formalism was developed.

A more systematic treatment of the problems discussed above will be presented in Chap. 2.

1.1.5 The Spin-Density Matrix and Its Basic Properties

1.1.5.1 Basic Definitions

Any question concerning the behavior of pure or mixed states can be answered by specifying the states, present in the mixture, and their statistical weights W_i . The actual calculations, however, are often very cumbersome. We will therefore now introduce an alternative method of characterizing pure and mixed states.

Consider a beam of N_a particles prepared in the state $|\chi_a\rangle$ and a second beam of N_b particles which have been prepared in the state $|\chi_b\rangle$ independently of the first one. In order to describe the joint beam we introduce an operator ρ by the expression.

$$\rho = W_a |\chi_a\rangle\langle\chi_a| + W_b |\chi_b\rangle\langle\chi_b| \quad (1.20)$$

with $W_a = N_a/N$, $W_b = N_b/N$, and $N = N_a + N_b$.

The operator ρ is called *density operator* or *statistical operator*. It describes the preparations which have been performed and, therefore, contains all the information obtained on the total beam. In this sense a mixture is completely specified by its density operator. In the special case of a pure state $|\chi\rangle$ the density operator is given by

$$\rho = |\chi\rangle\langle\chi| \quad (1.21a)$$

It will be seen later that it is usually more convenient to write ρ in matrix form. To this end we choose a set of basis states, commonly $|+1/2\rangle$ and $|-1/2\rangle$ and expand $|\chi_a\rangle$ and $|\chi_b\rangle$ in terms of this set according to (1.2):

$$\begin{aligned} |\chi_a\rangle &= a_1^{(a)} \left| +\frac{1}{2} \right\rangle + a_2^{(a)} \left| -\frac{1}{2} \right\rangle \\ |\chi_b\rangle &= a_1^{(b)} \left| +\frac{1}{2} \right\rangle + a_2^{(b)} \left| -\frac{1}{2} \right\rangle \end{aligned} \quad (1.22)$$

In the representation (1.1) we write

$$|\chi_a\rangle = \begin{pmatrix} a_1^{(a)} \\ a_2^{(a)} \end{pmatrix}, \quad |\chi_b\rangle = \begin{pmatrix} a_1^{(b)} \\ a_2^{(b)} \end{pmatrix} \quad (1.23a)$$

and for the adjoint states

$$\langle \chi_a | = \left(a_1^{(a)*}, a_2^{(a)*} \right), \quad \langle \chi_b | = \left(a_1^{(b)*}, a_2^{(b)*} \right) \quad (1.23b)$$

Applying the rules of matrix multiplication we obtain for the “outer product,” $|\chi_a\rangle\langle\chi_a|$:

$$\begin{aligned} |\chi_a\rangle\langle\chi_a| &= \begin{pmatrix} a_1^{(a)} \\ a_2^{(a)} \end{pmatrix} \begin{pmatrix} a_1^{(a)*} & a_2^{(a)*} \end{pmatrix} \\ &= \begin{pmatrix} |a_1^{(a)}|^2 & a_1^{(a)} a_2^{(a)*} \\ a_1^{(a)*} a_2^{(a)} & |a_2^{(a)}|^2 \end{pmatrix} \end{aligned} \quad (1.24)$$

and similarly for the product $|\chi_b\rangle\langle\chi_b|$. Substitution of these expressions into (1.20) yields the *density matrix*.

$$\rho = \begin{pmatrix} W_a |a_1^{(a)}|^2 + W_b |a_1^{(b)}|^2 & W_a a_1^{(a)} a_2^{(a)*} + W_b a_1^{(b)} a_2^{(b)*} \\ W_a a_1^{(a)*} a_2^{(a)} + W_b a_1^{(b)*} a_2^{(b)*} & W_a |a_1^{(a)}|^2 + W_b |a_2^{(b)}|^2 \end{pmatrix} \quad (1.25)$$

Since the basis states $|\pm 1/2\rangle$ have been used in deriving (1.25) this is said to be the *density matrix in the $\{|\pm 1/2\rangle\}$ representation*.

In order to make subsequent formulas more compact we define $|+1/2\rangle = |\chi_1\rangle$ and $|-1/2\rangle = |\chi_2\rangle$. In this notation the general element of the density matrix corresponding to the i th row and j th column is given by the expression

$$\langle \chi_i | \rho | \chi_j \rangle = W_a a_1^{(a)} a_j^{(a)*} + W_b a_i^{(b)} a_j^{(b)*} \quad (1.26)$$

with $i, j = 1, 2$.

Clearly the density matrix has a different form in different representations, whereas the operator (1.20) is independent of the choice of the basis states. It will always be assumed that the basis states are orthonormal, that is

$$\langle \chi_i | \chi_j \rangle = \delta_{ij} \quad (1.27)$$

where δ_{ij} denotes the Kronecker symbol and for $i = j$ condition (1.4) is satisfied.

In the normalization (1.4) the trace of the density matrix is given by

$$\text{tr } \rho = W_a + W_b = 1 \quad (1.28)$$

which is independent of the choice of the representation.

As an example, consider the case of a mixture consisting of N_1 particles which have been prepared in the state $|\chi_1\rangle = |+1/2\rangle$ and N_2 particles prepared

independently in the state $|\chi_2\rangle = | - 1/2\rangle$. The total beam is then represented by the density operator

$$\rho = W_1 \left| +\frac{1}{2} \right\rangle \left\langle +\frac{1}{2} \right| + W_2 \left| -\frac{1}{2} \right\rangle \left\langle -\frac{1}{2} \right| \quad (1.29a)$$

($W_i = N_i/N$) and the density matrix in the $\{|\pm 1/2\rangle\}$ representation is diagonal:

$$\langle \chi_i | \rho | \chi_j \rangle = W_i \delta_{ij} \quad (1.29b)$$

1.1.5.2 Significance of the Density Matrix

The diagonal elements of the density matrix

$$\langle \chi_i | \rho | \chi_i \rangle = W_a \left| a_i^{(a)} \right|^2 + W_b \left| a_i^{(b)} \right|^2 \quad (i = 1, 2) \quad (1.30)$$

have a direct physical meaning. Since the probability of finding a particle of the mixture in the state $|\chi_a\rangle$ is W_a and since the probability that $|\chi_a\rangle$ is in the state $|\chi_i\rangle$ is $\left| a_i^{(a)} \right|^2$, the product $W_a \left| a_i^{(a)} \right|^2$ is the probability that a particle originally prepared in the state $|\chi_a\rangle$ will be found in the state $|\chi_i\rangle$ after a measurement has been made. *The diagonal element (1.30) therefore gives the total probability of finding a particle in the corresponding basis state $|\chi_i\rangle$.*

Thus, if a beam described by a density operator ρ is sent through a Stern–Gerlach filter oriented parallel (antiparallel) to the z axis then the diagonal element $\langle \chi_1 | \rho | \chi_1 \rangle = \langle +\frac{1}{2} | \rho | +\frac{1}{2} \rangle = \langle (-\frac{1}{2} | \rho | -\frac{1}{2}) \rangle$ of ρ in the $\{|\pm \frac{1}{2}\rangle\}$ representation gives the probability that a particle will pass through the filter.

This result may be generalized to arbitrary states $|\chi\rangle$. Consider the matrix element $\langle \chi | \rho | \chi \rangle$ obtained by “sandwiching” the operator (1.20) between the state $|\chi\rangle$ and its adjoint $\langle \chi|$:

$$\begin{aligned} \langle \chi | \rho | \chi \rangle &= W_a \langle \chi | \chi_a \rangle \langle \chi_a | \chi \rangle + W_b \langle \chi | \chi_b \rangle \langle \chi_b | \chi \rangle \\ &= W_a |a^{(a)}|^2 + W_b |a^{(b)}|^2 \end{aligned} \quad (1.31)$$

where $a^{(a)} = \langle \chi_a | \chi \rangle$ and $a^{(b)} = \langle \chi_b | \chi \rangle$. Comparing (1.30) and (1.31) it can be seen that the *matrix element $\langle \chi | \rho | \chi \rangle$ is the total probability of finding a particle in the pure state $|\chi\rangle$ within a mixture which is represented by ρ* . That is, if a beam represented by ρ passes through a filter which only fully admits a beam in the state $|\chi\rangle$ then (1.31) gives the probability that any given particle of the beam will pass through the filter.

For example, suppose that a beam represented by the density matrix (1.29) is sent through a filter oriented in the y direction. The probability that a particle of the beam will pass through is then given by the matrix element $\langle +1/2, y | \rho | +1/2, y \rangle$.

Expressing $|+\frac{1}{2}, y\rangle$ in the $\{| \pm \frac{1}{2}\rangle$ representation [(1.13c)] and using (1.29) give

$$\begin{aligned} \left\langle +\frac{1}{2}, y \left| \rho \right| +\frac{1}{2}, y \right\rangle &= \frac{1}{2}(1, -i) \begin{pmatrix} W_1 & 0 \\ 0 & W_2 \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} \\ &= \frac{1}{2}(W_1 + W_2) \end{aligned}$$

The important point is that all the information on the spin state of any given beam can be obtained (in principle at least) by sending the beam through various Stern–Gerlach filters with different orientations. Consequently, once ρ is known, we can calculate the result of any such experiment by means of (1.31). In this sense, ρ contains all significant information on the spin state of a given beam.

1.1.5.3 The Number of Independent Parameters

We will now consider how many parameters are required in order to completely represent a given density matrix. A complex 2×2 matrix such as (1.25) has four complex elements $\langle \chi_i | \rho | \chi_j \rangle$ corresponding to eight real parameters. The density matrix is *Hermitian*; that is, ρ satisfies

$$\langle \chi_i | \rho | \chi_j \rangle = \langle \chi_j | \rho | \chi_i \rangle^* \quad (1.32)$$

This can be seen immediately from (1.25) or (1.26). Consequently, the diagonal elements are real, and the real and imaginary parts of the off-diagonal elements are related by the expressions

$$\begin{aligned} \operatorname{Re} \left\langle +\frac{1}{2} | \rho | -\frac{1}{2} \right\rangle &= \operatorname{Re} \left\langle -\frac{1}{2} | \rho | +\frac{1}{2} \right\rangle \\ \operatorname{Im} \left\langle +\frac{1}{2} | \rho | -\frac{1}{2} \right\rangle &= -\operatorname{Im} \left\langle -\frac{1}{2} | \rho | +\frac{1}{2} \right\rangle \end{aligned}$$

These relations reduce the number of independent real parameters to four. The normalization condition (1.28) fixes one further parameter so that the density matrix is completely characterized in terms of three real parameters. It follows from this that *three independent measurements must be performed in order to completely specify the density matrix for any given beam of spin-1/2 particles.*

It will be instructive to consider this result from another point of view. In (1.20) a density operator is defined from a knowledge of how a given beam has been prepared. This definition can be generalized for the case of any number of constituent beams. In order to write down the density operator or the corresponding density matrix from (1.20) and (1.25), it follows that all the pure states $|\chi_a\rangle, |\chi_b\rangle, \dots$ present in the mixture must be specified together with their statistical weights W_a, W_b, \dots . Only three real parameters, however, are required to completely characterize the density matrix of a beam of any complexity, as has been shown.

This is not as surprising as it may seem at first, since one and the same density matrix can represent many different mixtures prepared in entirely different ways. For example, consider a mixture represented by the density operator

$$\rho = \frac{1}{2} \left| +\frac{1}{2} \right\rangle \left\langle +\frac{1}{2} \right| + \frac{1}{2} \left| -\frac{1}{2} \right\rangle \left\langle -\frac{1}{2} \right|$$

and a mixture specified by the operator

$$\rho = \frac{1}{6} \left| +\frac{1}{2} \right\rangle \left\langle +\frac{1}{2} \right| + \frac{1}{6} \left| -\frac{1}{2} \right\rangle \left\langle -\frac{1}{2} \right| + \frac{1}{3} \left| +\frac{1}{2}, x \right\rangle \left\langle +\frac{1}{2}, x \right| + \frac{1}{3} \left| -\frac{1}{2}, x \right\rangle \left\langle -\frac{1}{2}, x \right|$$

By constructing the corresponding density matrices in the $\{|\pm 1/2\rangle\}$ representation and applying (1.13a), (1.13b) it can be shown that both beams are represented by the same density matrix:

$$\rho = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

It follows from (1.31) that the two beams will behave identically in all experiments with respect to their polarization properties. Conversely, a knowledge of the density matrix elements alone is insufficient to determine the method by which the beams have been prepared. In fact such information is insignificant. The only significant information is contained in the three independent parameters specifying the density matrix since these are sufficient to calculate the behavior of the corresponding beam in any polarization experiment. *For this reason we will henceforth consider two beams to be identical if they are described by the same density matrix.*

The definition (1.20) is usually of little importance, and instead of defining the density *operator* by specifying the constituent subbeams and their statistical weights we will apply a more operational point of view and define the density *matrix* by the results of three independent measurements. In the following section we will show how this can be achieved in a simple way by using the polarization vector.

1.1.5.4 Parametrization of the Density Matrix

If (1.20) is multiplied by the Pauli matrix σ_i the trace can be calculated as

$$\begin{aligned} \text{tr } \rho \sigma_i &= W_a \text{tr}(|\chi_a\rangle\langle\chi_a|\sigma_i) + W_b \text{tr}(|\chi_b\rangle\langle\chi_b|\sigma_i) \\ &= W_a \langle\chi_a|\sigma_i|\chi_a\rangle + W_b \langle\chi_b|\sigma_i|\chi_b\rangle \end{aligned} \quad (1.33)$$

This result can be obtained by using the explicit matrix representations (1.7) and (1.24) or, more directly, by applying the relation

$$\text{tr}(|\chi\rangle\langle\chi|\sigma_i) = \langle\chi|\sigma_i|\chi\rangle \quad (1.34)$$

Substituting (1.16a) into (1.33) gives the important result

$$\bullet \quad \text{tr } \rho \sigma_i = P_i \quad (1.35)$$

where P_i is the i th component of the polarization vector of the total beam.

Using this result the elements of ρ can be expressed in terms of the components P_i . By direct matrix manipulations it may be shown that in the $\{|\pm 1/2\rangle\}$ representation ρ is given by

$$\bullet \quad \rho = \frac{1}{2} \begin{pmatrix} 1 + P_z & P_x - iP_y \\ P_x + iP_y & 1 - P_z \end{pmatrix} \quad (1.36)$$

A more elegant method of deriving this result will be given in Sect. 1.1.6.

The three components P_x , P_y , P_z represent a minimum set of data which are required to specify the density matrix of any given beam and *we will henceforth regard the density matrix as defined by (1.36)*.

As an illustration of the use of (1.36) suppose a beam of particles, characterized by the matrix (1.36), passes a filter oriented in the z direction. The probability that a particle is admitted by the filter is given, according to (1.31), by the expression

$$\left\langle +\frac{1}{2} | \rho | +\frac{1}{2} \right\rangle = \frac{1}{2}(1 + P_z)$$

Similarly, applying (1.13a), (1.13c) and (1.36) the probabilities that a particle will pass through a filter oriented in the x and y directions are found to be

$$\begin{aligned} \left\langle +\frac{1}{2}, x | \rho | +\frac{1}{2}, x \right\rangle &= \frac{1}{2}(1 + P_x) \\ \left\langle +\frac{1}{2}, y | \rho | +\frac{1}{2}, y \right\rangle &= \frac{1}{2}(1 + P_y) \end{aligned}$$

Finally, we will give another useful representation of ρ obtained by transforming to a coordinate system x' , y' , z' , where z' is parallel to \mathbf{P} and x' and y' are chosen arbitrarily but orthogonal to each other and to z' . In this case, $P_{x'} = P_{y'} = 0$, $P_{z'} = |\mathbf{P}|$. As a result, in the representation with z' as the quantization axis, ρ is given by

$$\rho = \frac{1}{2} \begin{pmatrix} 1 + |\mathbf{P}| & 0 \\ 0 & 1 - |\mathbf{P}| \end{pmatrix} \quad (1.37a)$$

or by

$$\rho = \frac{1}{2}(1 - |\mathbf{P}|) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + |\mathbf{P}| \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad (1.37b)$$

If the beam under consideration is completely polarized, $|\mathbf{P}| = 1$, and

$$\rho = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad (1.38)$$

and the beam is in the pure state $|+1/2, z'\rangle$. If the beam is unpolarized $|\mathbf{P}| = 0$, and the corresponding density matrix is given by

$$\rho = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (1.39)$$

1.1.5.5 Identification of Pure States

In Sect. 1.1.2 it has been shown that a given beam is in a pure state if and only if its polarization vector has the maximum possible value of $|\mathbf{P}| = 1$. This result will now be put into a different form which is more useful for treating more complex systems.

Using (1.36) it can be shown that the trace of ρ^2 is given by

$$\begin{aligned} \text{tr}(\rho)^2 &= (1/2)(1 + P_x^2 + P_y^2 + P_z^2) \\ &= (1/2)(1 + |\mathbf{P}|^2) \end{aligned}$$

It follows from this that

$$\text{tr}(\rho)^2 = 1 \quad (1.40)$$

is a necessary and sufficient condition that the beam under consideration is in a pure state. [Note, the fact that the trace is equal to unity in (1.40) is a consequence of the normalization (1.28).]

In the case of a pure state the condition (1.40) gives an additional restriction on the density matrix elements. Thus *a pure state is characterized by two independent parameters only in accordance with (1.10).*

1.1.6 The Algebra of the Pauli Matrices

The discussion in Sect. 1.1.5 have shown that the result of any experiment performed with a given beam can be calculated from a knowledge of the corresponding density matrix. So far the required mathematical operations have to be carried out using a particular representation and applying the rules of matrix algebra. In general, this is a laborious and time-consuming procedure. In this section a more elegant method of performing the relevant calculations will be described.

The discussion will be based on the following fundamental relation between the Pauli matrices ($i, j = x, y, z$):

$$\sigma_i \sigma_j = \delta_{ij} \mathbf{1} + i \sum_k \epsilon_{ijk} \sigma_k \quad (1.41)$$

where δ_{ij} is the Kronecker symbol, $\mathbf{1}$ denotes the two-dimensional unit matrix, and

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } i, j, k \text{ is an even permutation of } XYZ \\ -1 & \text{if } i, j, k \text{ is an odd permutation of } XYZ \\ 0 & \text{if two of the indices are the same} \end{cases} \quad (1.42)$$

For example, for $i = j$ (1.41) becomes

$$\sigma_i^2 = 1 \quad (1.43a)$$

and for $i = x, j = y$:

$$\sigma_x \sigma_y = i \sigma_z, \quad \sigma_y \sigma_x = -i \sigma_z \quad (1.43b)$$

From (1.40) and (1.43) it follows that for $i \neq j$

$$\sigma_i \sigma_j + \sigma_j \sigma_i = 0 \quad (1.43c)$$

Equation 1.41 specifies completely the algebra of the Pauli matrices. Proofs of (1.41) can be found in any textbook on quantum mechanics.

The important property of (1.41) is that it reduces quadratic combinations of Pauli matrices to linear ones. This allows the calculation of traces of products of matrices σ_i by a stepwise reduction of the number of matrices occurring in the given trace. We give some examples. First, from (1.7) it can be seen that

$$\text{tr } \sigma_i = 0 \quad (1.44)$$

By taking the trace of (1.41) and using (1.44) it follows that

$$\text{tr } \sigma_i \sigma_j = 2\delta_{ij} \quad (1.45a)$$

A product of three Pauli matrices may first be reduced to a quadratic combination by means of (1.41):

$$\sigma_i \sigma_j \sigma_m = \delta_{ij} \sigma_m + i \sum_k \epsilon_{ijk} \sigma_k \sigma_m$$

Taking the trace of this expression and applying (1.44) and (1.45a) give

$$\text{tr } \sigma_i \sigma_j \sigma_m = 2i \sum_k \epsilon_{ijk} \delta_{km} = 2i \epsilon_{ijm} \quad (1.45b)$$

A further important property of the Pauli matrices is that any two-dimensional Hermitian matrix can be expressed as a linear combination of the unit matrix $\mathbf{1}$ and the matrices σ_i . For example, consider the density matrix. We make the “ansatz”

$$\rho = a\mathbf{1} + \sum_i b_i \sigma_i \quad (1.46)$$

In (1.46) the four coefficients a , b_x , b_y , b_z are unknowns which must be determined. Such an ansatz is possible because the hermiticity condition reduces the number of independent parameters determining ρ to four and there are four parameters in (1.46). One of the parameters can immediately be determined from the normalization condition (1.28), which gives, with the help of (1.44),

$$a = 1/2 \quad (1.47a)$$

Multiplying (1.46) by σ_j , taking the trace of the obtained expression, and using (1.44) and (1.45) give

$$\text{tr } \rho \sigma_j = 2 \sum_i b_i \delta_{ij} = 2b_j$$

The trace of ρ and σ_j gives the corresponding component of the polarization vector and from this it follows that

$$b_j = (1/2)P_j \quad (1.47b)$$

Inserting the results (1.47) into the ansatz (1.46) results in the expression

$$\bullet \quad \rho = \frac{1}{2} \left(\mathbf{1} + \sum_i P_i \sigma_i \right) \quad (1.48)$$

If the Pauli matrices are expressed in the form (1.7) then P can be obtained in the form (1.36). In the case of a pure state $|\chi\rangle$ characterized by

$$\rho^{(x)} = |\chi\rangle\langle\chi|$$

and denoting the polarization vector of the state $|\chi\rangle$ by $P^{(x)}$, we write

$$|\chi\rangle\langle\chi| = \frac{1}{2} \left(\mathbf{1} + \sum_i P_i^{(x)} \sigma_i \right) \quad (1.49)$$

This expression allows a simple determination of the probability $|\chi|\rho|\chi\rangle$. Equation 1.34 implies

$$\langle\chi|\rho|\chi\rangle = \text{tr}|\chi\rangle\langle\chi|\rho$$

Hence, using this result on the right-hand side of (1.49) gives

$$\begin{aligned}
\langle \chi | \rho | \chi \rangle &= \frac{1}{4} \operatorname{tr} \left[\left(1 + \sum_i P_i^{(x)} \sigma_i \right) \left(1 + \sum_j P_j \sigma_j \right) \right] \\
&= \frac{1}{4} \operatorname{tr} \left(1 + \sum_i P_i^{(x)} \sigma_i + \sum_j P_j \sigma_j + \sum_{ij} P_i^{(x)} P_j \sigma_i \sigma_j \right) \\
&= \frac{1}{2} (1 + \mathbf{P}^{(x)} \cdot \mathbf{P})
\end{aligned} \tag{1.50}$$

This result can be interpreted in the following way. A beam of particles may be characterized by a density matrix ρ . This beam may be passed through a Stern–Gerlach filter in a fixed orientation which only completely transmits a beam in the pure state $|\chi\rangle$ (that is, the filter is oriented parallel to $\mathbf{P}^{(x)}$).

The probability that a particle of the given beam will pass through the filter is then determined by the scalar product $\mathbf{P}^{(x)} \cdot \mathbf{P}$ of the two polarization vectors. The probability of transmission is a maximum if \mathbf{P} points in the direction of the magnetic field and is a minimum if \mathbf{P} is antiparallel to the filter direction. In particular, if the beam is unpolarized, then for *any* filter

$$\langle \chi | \rho | \chi \rangle = 1/2 \tag{1.51}$$

The derivation of (1.50) may serve as a first example of how calculations can be simplified by using (1.48) and the algebraic properties of the Pauli matrices.

1.1.7 Summary

The results obtained in the previous two sections allow a redefinition of the basic concepts used so far. We consider as the initial information on a given beam the values of the three components P_x , P_y , P_z of the polarization vector. \mathbf{P} can be determined, for example, by suitably chosen scattering experiments (for a detailed discussion of such experiments we refer particularly to Kessler 1976). When the polarization vector is known the density matrix can be obtained by means of (1.36) or (1.48). These expressions contain all information on the beam in condensed form. Their usefulness, particularly (1.48), in actual calculations will become evident in Sect. 2.5.

If $|\mathbf{P}| = 1$ the beam is said to be in a pure spin state, or, alternatively, all particles are in identical states. This joint state of all particles in the given beam is represented by assigning a single state vector to the whole beam. In this case two parameters are sufficient for a complete description of the spin state, for example, the polar angles θ and δ of \mathbf{P} , from which the corresponding state vector can be constructed by means of (1.10).

If $|\mathbf{P}| < 1$ the beam is said to be in a mixed state. Such states are characterized by three parameters, for example, the magnitude and the polar angles of \mathbf{P} .

1.2 Polarization States and Density Matrix of Photons

1.2.1 The Classical Concept of Wave Polarization

In this section, a description of photon polarization will be given. We will follow the arguments of Sect. 1.1 in order to become more familiar with the abstract concepts introduced there. We will begin with a brief account of the description of light polarization in classical optics.

A monochromatic electromagnetic wave is characterized by three quantities: its angular frequency ω , its wave vector $\mathbf{k} = (2\pi/\Lambda)\mathbf{n}$ (where \mathbf{n} is a unit vector in the direction of motion and Λ is the wavelength), and its state of polarization, which is defined by the vibrations of the electric field vector \mathbf{E} . The field vector \mathbf{E} of a monochromatic wave can be written in the form

$$\mathbf{E} = A\mathbf{e}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \quad (1.52)$$

where A is the amplitude and \mathbf{e} is the polarization vector. Because of the transverse nature of electromagnetic waves, \mathbf{e} is perpendicular to \mathbf{n} . In this section we will use a coordinate system x, y, z with the z axis parallel to \mathbf{n} , and restrict ourselves to a discussion of the polarization properties of light only. If \mathbf{E} vibrates along the x axis then the light is said to be linearly polarized along the x axis. The polarization vector is parallel to x and denoted by \mathbf{e}_x . If the electric vector oscillates along the y axis then the polarization is characterized by assigning a polarization vector \mathbf{e}_y to the beam pointing in the y' direction. A general polarization vector \mathbf{e} can always be expanded in terms of two orthogonal vectors, for example, \mathbf{e}_x and \mathbf{e}_y :

$$\mathbf{e} = a_1\mathbf{e}_x + a_2e^{i\delta}\mathbf{e}_y \quad (1.53)$$

where a_1 and a_2 are real coefficients. We will normalize (1.53) such that \mathbf{e} is always a unit vector in the sense that the scalar product of \mathbf{e} and its complex conjugate, \mathbf{e}^* , is equal to 1: $\mathbf{e} \cdot \mathbf{e}^* = 1$. The normalization condition is therefore

$$a_1^2 + a_2^2 = 1 \quad (1.54)$$

Equation 1.53 corresponds to a linear superposition of two waves of equal frequency and the same wave vector with amplitudes A_1 and A_2 , polarized along the x and y directions, respectively, with a definite phase difference δ :

$$\begin{aligned}\mathbf{E} &= A_1 \mathbf{e}_x e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} + A_2 \mathbf{e}_y e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t + \delta)} \\ &= A(a_1 \mathbf{e}_x + a_2 e^{i\delta} \mathbf{e}_y) e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}\end{aligned}$$

where a_1 and a_2 are the relative amplitudes of the waves normalized to unity: $a_i = A_i/A$ ($i = 1, 2$) with $A = (A_1^2 + A_2^2)^{1/2}$.

We can define a parameter β such that

$$a_1 = \cos \beta, \quad a_2 = \sin \beta \quad (1.55a)$$

[(1.54) is then automatically satisfied] and write the general polarization vector (1.53) in the form

$$\bullet \quad \mathbf{e} = \cos \beta \mathbf{e}_x + e^{i\delta} \sin \beta \mathbf{e}_y \quad (1.55b)$$

In order to become familiar with the use of this expression we will consider some specific cases.

1. Consider a superposition of two waves oscillating in phase, with relative amplitudes a_1 and a_2 and polarized along the x and y axes, respectively. From the relative amplitudes the parameter β can be determined, and inserting $\delta = 0$ into (1.55) the polarization vector of the resulting wave is found to be

$$\mathbf{e} = \cos \beta \mathbf{e}_x + \sin \beta \mathbf{e}_y \quad (1.56)$$

In this case it is possible to give a simple interpretation of β : \mathbf{e} is a real vector in the $x - y$ plane and (1.56) represents its decomposition in terms of the two orthogonal basis vectors \mathbf{e}_x and \mathbf{e}_y ; hence, β is the angle between \mathbf{e} and the x axis (Fig. 1.4).

2. A superposition of two waves with equal amplitudes $a_1 = a_2$ and a phase difference $\delta = \pm 90^\circ$ gives a wave with polarization vector

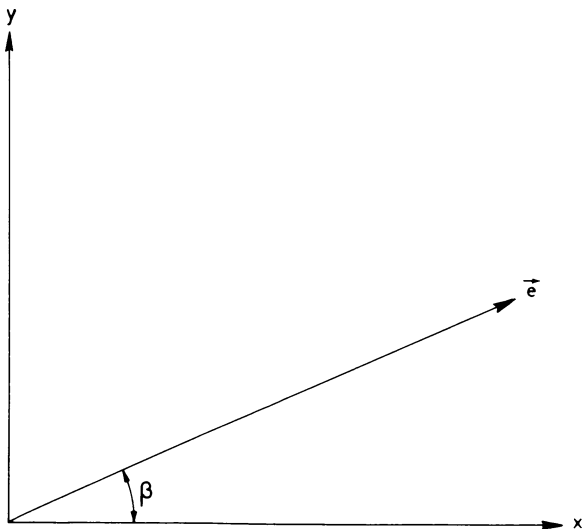
$$\mathbf{e} \sim \mathbf{e}_x \pm i \mathbf{e}_y$$

corresponding to left- and right-handed circular polarization. (For a further discussion of circular polarization see Sect. 1.2.3.)

3. If $a_1 \neq a_2$ and $\delta \neq 0$ we have the general case of elliptical polarization. In the following we will refer to a light wave as *completely polarized* if its polarization properties can be specified in terms of a single polarization vector \mathbf{e} [as in the case of the plane wave (1.1), for example]. It will be useful to reinterpret this definition in terms of some idealized experiments.

Following a treatment similar to that given in Sect. 1.1 the polarization properties of light can be discussed with the help of experiments with various optical polarization filters. We will assume that the filters used are always ideal in the

Fig. 1.4 Polarization vector of linearly polarized light



sense that the filter is completely transparent to light of a particular polarization and completely absorbs light of the opposite polarization. Hence light passing through the filter will emerge in a definite state of polarization. For example, a beam of light may pass a Nicol prism with its axis of transmission parallel to the x axis. The transmitted light is then linearly polarized along the x direction. Similarly, a beam of light passing a Nicol prism oriented parallel to an axis \mathbf{n} will emerge linearly polarized along this direction. If β is the angle between \mathbf{n} and the x axis the corresponding polarization vector is given by (1.56). Conversely, if linearly polarized light with polarization vector \mathbf{e} is passed through a Nicol prism it is always possible to find such an orientation of the prism which allows the whole beam to be transmitted. This occurs when the axis of transmission is parallel to \mathbf{e} . A circularly polarized wave will only be completely accepted by a circular polarization filter (for example, a suitably oriented series combination of a quarter wave plate and a Nicol prism).

By applying the converse of these arguments it can be seen that a light beam is completely polarized if such a filter can be found which completely admits the beam.

As is well known from optics light is usually not completely polarized. An ordinary light source consists of a large number of excited atoms each of which emits a pulse of light in a time of order $\sim 10^{-8}$ s independently of all other atoms. Because constantly new pulses will be contributing to the beam the overall polarization will change very rapidly and there will be no definite polarization vector which is characteristic of the total beam. In the following sections we will consider the problem of describing beams of this kind.

1.2.2 Pure and Mixed Polarization States of Photons

When the theory of relativistic quantum mechanics is applied to the electromagnetic field it follows that in interaction with matter the wave behaves as if it were composed of photons. We will start our discussion with the following definition:

- A beam of photons is said to be in a *pure polarization state* if the beam is completely polarized in the sense explained in Sect. 1.2.1.

In terms of our idealized experiments this definition may be reinterpreted as follows: if it is possible to find such a filter which completely admits a beam of photons then the beam is said to be in a pure polarization state. Alternatively we may say that all photons of the beam can be considered to be in one and the same polarization state. This joint state of *all* the photons can be described in terms of a *single* state vector which we will denote by $|e\rangle$, by which is meant the polarization state of any photon in the beam which classically has polarization vector \mathbf{e} . For example, the state vectors $|e_x\rangle$ and $|e_y\rangle$ denote the polarization state of photons which are completely transmitted by a Nicol prism oriented in the x and y directions, respectively.

The states $|e_x\rangle$ and $|e_y\rangle$ can be taken as basis states and any state $|e\rangle$ can be written as a linear superposition:

$$|e\rangle = a_1|e_x\rangle + a_2|e_y\rangle \quad (1.57)$$

or

$$|e\rangle = \cos \beta |e_x\rangle + e^{i\delta} \sin \beta |e_y\rangle \quad (1.58)$$

These equations are exactly analogous to (1.53) and (1.55), respectively.

These considerations are similar to the discussions in Sect. 1.1. All the experiments and results for spin-1/2 particles and Stern–Gerlach filters which have been described previously can be repeated with photons and polarization filters. In particular, it should be noted that a_1^2 and a_2^2 are the probabilities that a photon in the polarization state (1.57) will pass through a Nicol prism oriented parallel to the x or y axis, respectively.

As shown by (1.57) and (1.58) any superposition of two (or more) states which have a definite phase δ necessarily results in a pure state. Thus, in order to describe light which is not completely polarized, it is necessary to consider superposition states which do not have a definite phase relation, that is, we have to introduce the concept of a mixture. In general, *a beam of photons is said to be in a mixed state or mixture if it is not possible to describe the beam in terms of a single state vector.*

It is useful to visualize the concept of a mixture in terms of some idealized experiments. Consider two light sources emitting *independently* of each other, where by “independently” is meant that there is no definite phase relationship between the two sources (that is, the relative phase changes much more rapidly than the observation time in an unpredictable manner). Both sources are provided with a polarization filter so that the first source emits a beam of intensity I_1 of definite polarization $|e_1\rangle$ and the second one a beam of intensity I_2 and polarization $|e_2\rangle$.

If the two beams are combined and the polarization properties of the total beam investigated by sending it through various filters, it will be found that, irrespective of the nature of the filter, the transmitted intensity is always less than the incident one. Thus, by definition, the total beam is in a mixed polarization state.

It is not possible to completely characterize a mixture by a single-state vector $|e\rangle$. In particular, the mixture cannot be represented as a linear superposition of the states $|e_1\rangle$ and $|e_2\rangle$. The reason for this is that, as has been discussed in Sects. 1.1.1–1.1.4, there is no definite phase δ between the constituent beams with which a definite state vector $|e\rangle$ can be constructed.

1.2.3 The Quantum Mechanical Concept of Photon Spin

In classical optics the polarization of light is explained in terms of the vibrations of the electric field vector. We will now investigate how polarization states can be interpreted in terms of the characteristic properties of photons.

To this end we will consider the possible spin states of photons. There are certain limitations to the concept of the photon spin. The total angular momentum \mathbf{J} of any particle is the resultant of its spin \mathbf{S} and its orbital angular momentum \mathbf{L} . Since the rest mass of a photon is zero, the usual definition of spin as the total angular momentum of a particle at rest is inapplicable for a photon. Strictly, only the total angular momentum \mathbf{J} of the photon has any physical meaning. However, it is convenient to define a spin and an orbital angular momentum in a formal sense. The photon spin is given the value 1 corresponding to the fact that the wave function is a vector [as shown, for example, by (1.52)]. The value of the orbital angular momentum is related to the multipoles which occur in the wave function (see, for example, Landau and Lifschitz 1965).

In general, if a spin-1 particle has a well-defined momentum \mathbf{p} the components of its spin along its direction of motion can take three values: $+1$, -1 , 0 . However, because of the transverse nature of electromagnetic waves the value 0 must be excluded for photons. The component of the photon spin along the direction of propagation \mathbf{n} , which we will denote by the symbol λ , can therefore only have the values $\lambda = +1$ (“spin up”) and $\lambda = -1$ (“spin down”).

It is important to note that the two photon states with spin up and spin down with respect to \mathbf{n} as quantization axis have direct physical meaning. Since the component of the orbital angular momentum vanishes in the direction of propagation \mathbf{n} we have $\mathbf{J} \cdot \mathbf{n} = (\mathbf{L} + \mathbf{S}) \cdot \mathbf{n} = \mathbf{S} \cdot \mathbf{n} = \lambda$; consequently,

- λ is the component of the total angular momentum of the photon in the direction of propagation \mathbf{n} .

The component of the spin in the direction of motion is generally called the helicity and we will refer to photon states with definite values $\lambda = \pm 1$ as *helicity states*.

Classically, when a circularly polarized light beam is directed at a target the electrons in the target are set into circular motion in response to the rotating electric field of the wave. This suggests that there is a relationship between circularly polarized light and photons in definite states of angular momentum.

In fact it has been shown in quantum electrodynamics *that photons with definite helicity are related to left-handed and right-handed circular polarization states*. Unfortunately, this notation is not unambiguous and we will adopt the following convention. We will denote the polarization vector and state of photons with helicity $\lambda = 1$ by \mathbf{e}_{+1} and $|+1\rangle$, respectively, and refer to light of positive helicity as *right-handed circularly polarized*.

Similarly, if $\lambda = -1$ we will denote the polarization vector and state of such light by \mathbf{e}_{-1} and $|-1\rangle$, respectively, and refer to the light as *left-handed circularly polarized*.

Note that in the terminology of classical optics the opposite convention is usually adopted: Light of positive (negative) helicity is called left-handed (right-handed) circularly polarized. We will always use the helicity state notation in order to avoid this ambiguity. With this convention the vectors \mathbf{e}_{+1} and \mathbf{e}_{-1} and the states $|\pm 1\rangle$ are then determined apart from a phase factor which has little significance and we write

$$\mathbf{e}_{\pm 1} = \mp \frac{1}{2^{1/2}}(\mathbf{e}_x \pm i \mathbf{e}_y) \quad (1.59)$$

for the polarization vector and for the corresponding states:

$$|\pm 1\rangle = \mp \frac{1}{2^{1/2}}(|e_x\rangle \pm i|e_y\rangle) \quad (1.60)$$

(see, for example, Messiah 1965).

In particular for problems where questions of angular momentum must be taken explicitly into account, it is convenient to use the helicity states as basis states instead of $|e_x\rangle$ and $|e_y\rangle$. We will therefore write the general polarization state $|e\rangle$ in the form

$$|e\rangle = a_1|+1\rangle + a_2|-1\rangle \quad (1.61)$$

There is a close formal analogy between photons and spin-1/2 particles. Because there are only two possible values of the helicity $\lambda = \pm 1$ (corresponding to states with spin up and spin down with respect to \mathbf{n} as the quantization axis) these states can be represented by two-dimensional column vectors *as long as \mathbf{n} is used as the axis of quantization* (z axis). The basis states can then be written in a similar way to (1.1):

$$|+1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |-1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (1.62)$$

In this representation the general pure state (1.61) is described by the row vector

$$|e\rangle = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad (1.63a)$$

and its adjoint by the column vector

$$|e\rangle = (a_1^*, a_2^*) \quad (1.63b)$$

For example, the state of light beams which are completely linearly polarized along the x and y axes, respectively, is obtained by inverting (1.60):

$$|e_x\rangle = -\frac{1}{2^{1/2}}(|+1\rangle - |-1\rangle) \quad (1.64a)$$

$$|e_y\rangle = -\frac{i}{2^{1/2}}(|+1\rangle + |-1\rangle) \quad (1.64b)$$

The interpretation of these linear superposition states is analogous to that given in Sect. 1.1.4.

As another example, consider a beam of photons prepared in the pure state $|e_x\rangle$. It can be seen from (1.64a) that these photons have no definite helicity. However, in any experiment performed on the beam, in which the angular momentum is actually measured, any photon in the beam will be forced into one of the angular momentum eigenstates, $|+1\rangle$ or $|-1\rangle$, with equal probability. In any such experiment any photon of the beam will therefore transfer a definite amount of angular momentum, either $\lambda = +1$ or $\lambda = -1$. Since the corresponding probabilities are equal, the net angular momentum, transferred by the total beam, is zero.

1.2.4 The Polarization Density Matrix

A compact expression of the polarization properties of photons is contained in the corresponding density matrix. In Sect. 1.2.5 we will give an operational definition of the photon density matrix. But in this section we will follow the arguments of Sect. 1.1.5.

Consider a beam of photons which is a mixture of two beams which have been prepared independently in the states $|e_a\rangle$ and $|e_b\rangle$, with intensities I_a and I_b , respectively. The density operator characterizing the total beam is defined by the expression

$$\rho' = W_a |e_a\rangle \langle e_a| + W_b |e_b\rangle \langle e_b| \quad (1.65)$$

with $W_a = I_a/I$ and $W_b = I_b/I$ and $I = I_a + I_b$. In order to obtain the density matrix it is necessary to choose a particular representation. We will use the helicity states as basis and expand the two states $|e_a\rangle$ and $|e_b\rangle$ according to (1.61) as

$$\begin{aligned} |e_a\rangle &= a_1^{(a)} \begin{vmatrix} +1 \end{vmatrix} + a_2^{(a)} \begin{vmatrix} -1 \end{vmatrix} \\ |e_b\rangle &= a_1^{(b)} \begin{vmatrix} +1 \end{vmatrix} + a_2^{(b)} \begin{vmatrix} -1 \end{vmatrix} \end{aligned}$$

Using the explicit representation (1.62) and applying the rule (1.24) the density matrix in the helicity representation is found to be

$$\rho' = \begin{pmatrix} W_a |a_1^{(a)}|^2 + W_b |a_1^{(b)}|^2 & W_a a_1^{(a)} a_2^{(a)*} + W_b a_1^{(b)} a_2^{(b)*} \\ W_a a_1^{(a)*} a_2^{(a)} + W_b a_1^{(b)*} a_2^{(b)} & W_a |a_2^{(a)}|^2 + W_b |a_2^{(b)}|^2 \end{pmatrix} \quad (1.66)$$

From the explicit representation (1.66) it follows that ρ' is normalized according to

$$\text{tr } \rho' = W_a + W_b = 1 \quad (1.67)$$

It is often more convenient to normalize ρ in such a way that its trace is equal to the total intensity of the corresponding photon beam. This can be achieved by substituting the intensities I_a and I_b for W_a and W_b in the definition (1.65) and (1.66). The density operator in this normalization is then given by

$$\rho = I_a |e_a\rangle\langle e_a| + I_b |e_b\rangle\langle e_b| \quad (1.68)$$

The trace of the density matrix is then

$$\text{tr } \rho = I_a + I_b = I \quad (1.69)$$

The density matrix ρ (and ρ') has the following properties (the proofs are like those in Sect. 1.1.5):

1. The diagonal elements $\langle +1 | \rho' | +1 \rangle$ and $\langle -1 | \rho' | -1 \rangle$ of the matrix (1.66) give the probability of finding a photon in the beam in the corresponding helicity state. In the normalization (1.69) the diagonal elements $\langle +1 | \rho | +1 \rangle$ and $\langle -1 | \rho | -1 \rangle$ give the corresponding intensities.
2. If the beam under consideration is sent through a filter which fully admits only photons in the pure state $|e\rangle$ then the element

$$\langle e | \rho' | e \rangle = W_a |a^{(a)}|^2 + W_b |a^{(b)}|^2 \quad (1.70)$$

gives the probability that a photon of the beam will be transmitted through the filter where we used the notation

$$a^{(a)} = \langle e | e_a \rangle$$

$$a^{(b)} = \langle e | e_b \rangle$$

The element $\langle e | \rho | e \rangle$ obtained from the operator (1.68) gives the transmitted intensity:

$$\langle e | \rho | e \rangle = I_a |a^{(a)}|^2 + I_b |a^{(b)}|^2 \quad (1.71)$$

Since any information on the polarization properties of a given beam can be obtained in principle by allowing the beam to pass through various polarization filters, the result of any such experiment can be calculated by using formulas (1.70) or (1.71). From this it can be concluded that all information on the polarization state of a given beam is contained in its density matrix.

3. The hermiticity condition (1.32) reduces the number of independent parameters to four. One of these is usually the total beam intensity I . If I is not of interest it can be dropped by normalizing as in (1.67) where ρ' then is specified by three real parameters similar to the density matrix of spin-1/2 particles.

It follows that four independent measurements are required in order to specify the matrix ρ completely for any given beam (one of these is the determination of the intensity I). If I is dropped the matrix (1.66), which requires a set of three independent measurements, is obtained. The result of any further experiment can then be calculated by applying (1.70) or (1.71).

4. In Chap. 2 we will prove that, in general, a necessary and sufficient condition that a given photon density matrix describes a pure state is given by

$$\text{tr}(\rho^2) = (\text{tr} \rho)^2 = I^2 \quad (1.72)$$

In the normalization (1.67) this reduces to (1.40):

$$\text{tr}(\rho^2) = 1$$

In general, the photon density matrix satisfies

$$\text{tr}(\rho^2) \leq I^2 \quad (1.73a)$$

1.2.5 Stokes Parameter Description

1.2.5.1 Parametrization of ρ in Terms of the Stokes Parameters

We will henceforth adopt the normalization (1.69). In the preceding section it has been seen that four independent measurements must be performed in order to completely determine the polarization state of any given beam. The most convenient set of measurements is that one which gives the following information:

1. The total intensity I of the beam;
2. The degree of linear polarization with respect to the x and y axes, defined as

$$\eta_3 = \frac{I(0) - I(90^\circ)}{I} \quad (1.74a)$$

where $I(\beta)$ denotes the intensity transmitted by a Nicol prism oriented at an angle β with respect to the x axis;

3. The degree of linear polarization with respect to two orthogonal axes oriented at 45° to the x axes

$$\eta_1 = \frac{I(45^\circ) - I(135^\circ)}{I} \quad (1.74b)$$

4. The degree of circular polarization defined as

$$\eta_2 = \frac{I_{+1} - I_{-1}}{I} \quad (1.74c)$$

where I_{+1} (and I_{-1}) are the intensities of light transmitted by polarization filters which fully transmit only photons with positive (negative) helicity.

The parameters 1–4 are called *Stokes parameters*. A detailed description of these can be found in Born and Wolf (1970). (See also McMaster 1954; Farago 1971.)

We will now relate the Stokes parameters to the elements of the density matrix. Denoting the elements of ρ by $\rho_{\lambda'\lambda} \equiv \langle \lambda' | \rho | \lambda \rangle$ we write

$$\rho = \begin{pmatrix} \rho_{+1,+1} & \rho_{+1,-1} \\ \rho_{-1,+1} & \rho_{-1,-1} \end{pmatrix} \quad (1.75)$$

where $\rho_{+1,-1} = \rho_{-1,+1}^*$ because of the hermicity condition (1.32). Applying (1.69) the total intensity is given by

$$I = \rho_{11} + \rho_{-1,-1} \quad (1.76a)$$

In order to obtain η_3 we have to calculate the intensities $I(0)$ and $I(90^\circ)$. From relation (1.71) these are given by

$$\begin{aligned} I(0) &= \langle e_x | \rho | e_x \rangle \\ I(90^\circ) &= \langle e_y | \rho | e_y \rangle \end{aligned}$$

In the helicity representation the state vectors $|e_x\rangle$ and $|e_y\rangle$ are expressed by (1.64); hence

$$\begin{aligned} I(0) &= (1/2)(-1, +1) \begin{pmatrix} \rho_{11} & \rho_{1,-1} \\ \rho_{-1,1} & \rho_{-1,-1} \end{pmatrix} \begin{pmatrix} -1 \\ +1 \end{pmatrix} \\ &= (1/2)(\rho_{11} - \rho_{1,-1} - \rho_{-1,1} + \rho_{-1,-1}) \end{aligned}$$

Similarly, we obtain

$$\begin{aligned} I(90^\circ) &= (1/2)(-i, -i) \begin{pmatrix} \rho_{11} & \rho_{1,-1} \\ \rho_{-1,1} & \rho_{-1,-1} \end{pmatrix} \begin{pmatrix} i \\ i \end{pmatrix} \\ &= (1/2)(\rho_{11} + \rho_{1,-1} + \rho_{-1,1} + \rho_{-1,-1}) \end{aligned}$$

It therefore follows that

$$I\eta_3 = -(\rho_{1,-1} + \rho_{-1,1}) \quad (1.76b)$$

In the same way we calculate the parameter $I\eta_1$, defined by (1.74c). In this case the axes of transmission of the Nicols are set at angles 45° and 135° to the x axes, respectively. The intensities transmitted by these prisms are then given by

$$\begin{aligned} I(45^\circ) &= \langle e_1 | \rho | e_1 \rangle \\ I(135^\circ) &= \langle e_2 | \rho | e_2 \rangle \end{aligned}$$

where $|e_1\rangle$ denotes a photon state which is fully transmitted by the first prism; that is,

$$|e_1\rangle = (1/2^{1/2})(|e_x\rangle + |e_y\rangle)$$

where (1.56) has been used with $\beta = 45^\circ$. Similarly, $|e_2\rangle$ is a photon state which is fully transmitted by the second prism and can be expressed in terms of $|e_x\rangle$ and $|e_y\rangle$ by inserting $\beta = 135^\circ$ in (1.56):

$$|e_2\rangle = (1/2^{1/2})(-|e_x\rangle + |e_y\rangle)$$

Transforming $|e_x\rangle$ and $|e_y\rangle$ to the helicity basis gives

$$I\eta_1 = -i(\rho_{1,-1} - \rho_{-1,1}) \quad (1.76c)$$

Similarly,

$$I\eta_2 = \rho_{11} - \rho_{-1,1} \quad (1.76d)$$

By inverting these equations the elements $\rho_{\lambda'\lambda}$ can be expressed in terms of the Stokes parameters:

$$\bullet \quad \rho = \frac{I}{2} \begin{pmatrix} 1 + \eta_2 & -\eta_3 + i\eta_1 \\ -\eta_3 - i\eta_1 & 1 - \eta_2 \end{pmatrix} \quad (1.77)$$

We will use this form of the density matrix throughout this book.

1.2.5.2 Examples

It follows from (1.58) that any pure polarization state can be parametrized in the form

$$|e\rangle = \cos \beta |e_x\rangle + e^{i\delta} \sin \beta |e_y\rangle \quad (1.78)$$

The corresponding density operator is given by $\rho = I|e\rangle\langle e|$. We will calculate the Stokes parameters characterizing a beam in the state (1.78). We have

$$I(0) = \langle e_x | \rho | e_x \rangle = I |\langle e_x | e \rangle|^2 = I \cos^2 \beta$$

$$I(90)^\circ = \langle e_y | \rho | e_y \rangle = I |\langle e_y | e \rangle|^2 = I \sin^2 \beta$$

from which follows

$$\eta_3 = \cos 2\beta \quad (1.79a)$$

Similarly we find

$$\eta_1 = \sin 2\beta \cos \delta \quad (1.79b)$$

$$\eta_2 = \sin 2\beta \sin \delta \quad (1.79c)$$

For example, the pure state $|e_x\rangle$, characterizing a beam of light with polarization vector pointing in the x direction, is obtained by inserting $\delta = 0$, $\beta = 0$ in (1.78). From (1.79) we obtain the Stokes parameters $\eta_3 = 1$, $\eta_1 = \eta_2 = 0$. Inserting these values into the density matrix (1.77) gives

$$\rho = \frac{I}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \quad (1.80a)$$

A beam which is linearly polarized in the y direction can be specified by the parameters $\beta = 90^\circ$, $\delta = 0$ so that

$$\eta_3 = -1, \quad \eta_1 = \eta_2 = 0$$

and

$$\rho = \frac{I}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (1.80b)$$

Similarly, as shown in Sect. 1.2.1 a beam linearly polarized in a direction with angle β with respect to the x axis is described by inserting $\delta = 0$ in (1.78) and (1.79). The Stokes parameters are therefore given by $\eta_3 = \cos 2\beta$, $\eta_1 = \sin 2\beta$, $\eta_2 = 0$, and the corresponding density matrix is

$$\rho = \frac{I}{2} \begin{pmatrix} 1 & -\cos 2\beta + i \sin 2\beta \\ -\cos 2\beta - i \sin 2\beta & 1 \end{pmatrix} \quad (1.80c)$$

Left- and right-handed polarized light are represented by the density matrices

$$\rho = I \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad (1.81a)$$

and

$$\rho = I \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad (1.81b)$$

respectively.

Once, the Stokes parameters, and hence the density matrix, have been determined, it is straightforward to derive a useful expression for the intensity, I_e , of light transmitted by a filter which only admits photons in the state $|e\rangle$: The required element is $\langle e|\rho|e\rangle$, which is

$$\bullet \quad I_e = (I/2)(1 + \eta_3 \cos 2\beta + \eta_1 \sin 2\beta \cos \delta + \eta_2 \sin 2\beta \sin \delta) \quad (1.82)$$

Note that the parameters β and δ describe the transmitted beam, whereas the incident beam is specified in terms of the Stokes parameters in (1.82).

1.2.5.3 Degree of Polarization

We will now introduce a further notation which will be useful in later discussions. From condition (1.73a) and (1.77) it follows that the Stokes parameters are restricted by the condition

$$\eta_1^2 + \eta_2^2 + \eta_3^2 \leq 1 \quad (1.83)$$

The equality sign holds only if the photons in the beam under discussion are in a pure polarization state. Alternatively, the beam is completely polarized (in the sense explained in Sect. 1.2.1) if and only if the relation

$$\eta_1^2 + \eta_2^2 + \eta_3^2 = 1 \quad (1.84a)$$

holds. These conditions may be conveniently expressed by introducing the quantity

$$P = (\eta_1^2 + \eta_2^2 + \eta_3^2)^{1/2} \quad (1.85)$$

It follows from (1.83) that P is restricted by

$$P \leq 1 \quad (1.86a)$$

Equations 1.83 and 1.85 can then be summarized as follows:

- A given beam of photons is in a pure polarization state if and only if $P = 1$. If $P < 1$ the beam is in a mixed state.

If a beam is such that $P > 0$ we will refer to the beam as *polarized* (*completely polarized* if $P = 1$); if $P = 0$ we will refer to the beam as *unpolarized*. In the latter case all Stokes parameters vanish and the corresponding density matrix is given by

$$\rho = \frac{I}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (1.87)$$

Since the Stokes parameters vanish in *any* representation when $P = 0$, (1.87) is independent of the choice of the basis states. Any mixture of independently prepared states $|e_1\rangle$ and $|e_2\rangle$ of opposite polarization (for example, $|+1\rangle$ and $|-1\rangle$ or $|e_x\rangle$ and $|e_y\rangle$) and equal intensities $I_1 = I_2 = I/2$ is represented by the density matrix (1.87). All these mixtures behave identically in their polarization properties and can be used as models for unpolarized light.

1.2.5.4 “Operational” Definition of ρ

At this point we will invert some of the results given above, as follows. In order to determine the polarization properties of a given light beam it is necessary to perform four independent measurements by, for convenience, determining the Stokes parameters. These four parameters then serve as data which enable the density matrix to be *defined* by (1.77). The result of any further experiment performed on the beam can then be calculated with the help of (1.71) or (1.82).

A beam of photons is in a pure polarization state if and only if $P = 1$ (completely polarized beam). In this case the polarization state of the beam can be represented by a single state vector $|e\rangle$. In this case the Stokes parameters are not independent; because of condition (1.84a) three of the parameters suffice for a complete characterization of the beam [two parameters in the case of the normalization (1.67)].

Finally, a beam of photons is in a mixed state if $P < 1$. In the special case $P = 0$ the beam is unpolarized and represented by the density matrix (1.87).



<http://www.springer.com/978-3-642-20560-6>

Density Matrix Theory and Applications

Blum, K.

2012, XVIII, 346 p., Hardcover

ISBN: 978-3-642-20560-6