3

Polarization: photons and spin-1/2 particles

In this chapter we build up the basic concepts of quantum mechanics using two simple examples, following a heuristic approach which is more inductive than deductive. We start with a familiar phenomenon, that of the polarization of light, which will allow us to introduce the necessary mathematical formalism. We show that the description of polarization leads naturally to the need for a two-dimensional complex vector space, and we establish the correspondence between a polarization state and a vector in this space, referred to as the space of polarization states. We then move on to the quantum description of photon polarization and illustrate the construction of probability amplitudes as scalar products in this space. The second example will be that of spin 1/2, where the space of states is again two-dimensional. We construct the most general states of spin 1/2 using rotational invariance. Finally, we introduce dynamics, which allows us to follow the time evolution of a state vector.

The analogy with the polarization of light will serve as a guide to constructing the quantum theory of photon polarization, but no such classical analog is available for constructing the quantum theory for spin 1/2. In this case the quantum theory will be constructed without reference to any classical theory, using an assumption about the dimension of the space of states and symmetry principles.

3.1 The polarization of light and photon polarization

3.1.1 The polarization of an electromagnetic wave

The polarization of light or, more generally, of an electromagnetic wave, is a familiar phenomenon related to the vector nature of the electromagnetic field. Let us consider a plane wave of monochromatic light of frequency ω propagating in the positive z direction. The electric field $\vec{E}(t)$ at a given point is a vector orthogonal to the direction of propagation. It therefore lies in the xOy plane and has components $\{E_x(t), E_y(t), E_z(t) = 0\}$ (Fig. 3.1). The most general case is that of elliptical polarization, where the electric field has the form

$$\vec{E}(t) = \begin{cases} E_x(t) = E_{0x}\cos(\omega t - \delta_x) \\ E_y(t) = E_{0y}\cos(\omega t - \delta_y) \end{cases}. \tag{3.1}$$

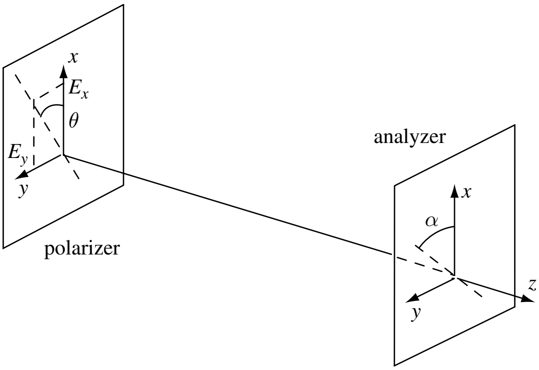

Fig. 3.1. A polarizer–analyzer ensemble.

We have not made the z dependence explicit because we are only interested in the field in a plane $z =$ constant. By a suitable choice of the origin of time, it is always possible to choose $\delta_x = 0,\ \delta_y = \delta$. The intensity $\mathcal{I}$ of the light wave is proportional to the square of the electric field:

$$\mathcal{I} = \mathcal{I}_x + \mathcal{I}_y = k(E_{0x}^2 + E_{0y}^2) = kE_0^2, \tag{3.2}$$

where k is a proportionality constant which need not be specified here. When $\delta = 0$ or π, the polarization is *linear*: if we take $E_{0x} = E_0 \cos\theta$, $E_{0y} = E_0 \sin\theta$, Eq. (3.1) for $\delta_x = \delta_y = 0$ shows that the electric field oscillates in the $\hat{n}_\theta$ direction of the xOy plane, making an angle θ with the Ox axis. Such a light wave can be obtained using a linear polarizer whose axis is parallel to $\hat{n}_\theta$.

When we are interested only in the polarization of this light wave, the relevant parameters are the ratios $E_{0x}/E_0 = \cos\theta$ and $E_{0y}/E_0 = \sin\theta$, where θ can be chosen to lie in the range $[0, \pi]$. Here E_0 is a simple proportionality factor which plays no role in the description of the polarization. We can establish a correspondence between waves linearly polarized in the Ox and Oy directions and orthogonal unit vectors $|x\rangle$ and $|y\rangle$ in the xOy plane forming an orthonormal basis in this plane. The most general state of *linear* polarization in the $\hat{n}_\theta$ direction will correspond to the vector $|\theta\rangle$ in the xOy plane:

$$|\theta\rangle = \cos\theta|x\rangle + \sin\theta|y\rangle, \tag{3.3}$$

which also has unit norm:

$$\langle\theta|\theta\rangle = \cos^2\theta + \sin^2\theta = 1.$$

The fundamental reason for using a vector space to describe polarization is the *superposition principle*: a polarization state can be decomposed into two (or more) other states, or, conversely, two polarization states can be added together vectorially. To illustrate decomposition, let us imagine that a wave polarized in the $\hat{n}_\theta$ direction passes through a second polarizer, called an analyzer, oriented in the $\hat{n}_\alpha$ direction of the xOy plane making an angle α with Ox (Fig. 3.1). Only the component of the electric field in the

$\hat{n}_\alpha$ direction, that is, the projection of the field on $\hat{n}_\alpha$, will be transmitted. The amplitude of the electric field will be multiplied by a factor $\cos(\theta - \alpha)$ and the light intensity at the exit from the analyzer will be reduced by a factor $\cos^2(\theta - \alpha)$. We shall use $\overline{a}(\theta \to \alpha)$ to denote the projection factor, which we refer to as the *amplitude of the $\hat{n}_\theta$ polarization in the $\hat{n}_\alpha$ direction*. This amplitude is just the scalar product of the vectors $|\theta\rangle$ and $|\alpha\rangle$:

$$\overline{a}(\theta \to \alpha) = \langle \alpha | \theta \rangle = \cos(\theta - \alpha) = \hat{n}_\alpha \cdot \hat{n}_\theta. \tag{3.4}$$

The intensity at the exit of the analyzer is given by the Malus law:

$$\mathcal{J} = \mathcal{J}_0 |\overline{a}(\theta \to \alpha)|^2 = \mathcal{J}_0 |\langle \alpha | \theta \rangle|^2 = \mathcal{J}_0 \cos^2(\theta - \alpha) \tag{3.5}$$

if $\mathcal{J}_0$ is the intensity at the exit of the polarizer. Another illustration of decomposition is given by the apparatus of Fig. 3.2. Using a uniaxial birefringent plate perpendicular to the direction of propagation and with optical axis lying in the xOz plane, a light beam can be decomposed into a wave polarized in the Ox direction and a wave polarized in the Oy direction. The wave polarized in the Ox direction propagates in the direction of the extraordinary ray refracted at the entrance and exit of the plate, and the wave polarized in the Oy direction follows the ordinary ray propagating in a straight line.

The addition of two polarization states can be illustrated using the apparatus of Fig. 3.3. The two beams are recombined by a second birefringent plate, symmetrically located relative to the first with respect to a vertical plane, before the beam passes through the analyzer.[1] In order to simplify the arguments, we shall neglect the phase difference

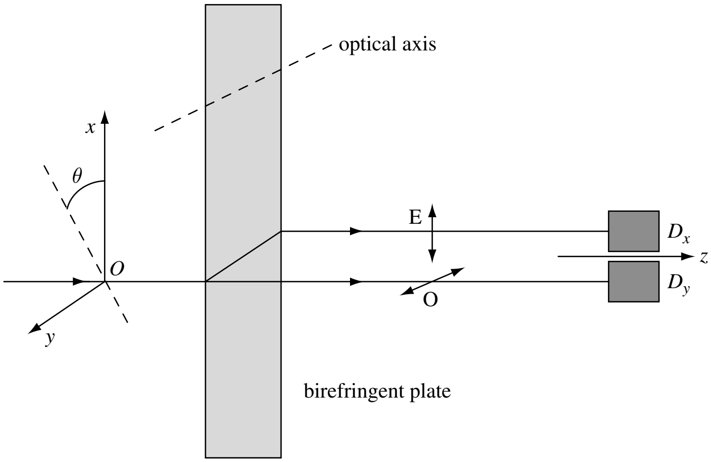

Fig. 3.2. Decomposition of the polarization by a birefringent plate. The ordinary ray O is polarized horizontally, and the extraordinary ray E is polarized vertically.

[1] This recombination of amplitudes is possible because two beams from the same source are coherent. Of course, it would be impossible to add the amplitudes of two polarized beams from different sources; the situation is identical to that in the case of interference.

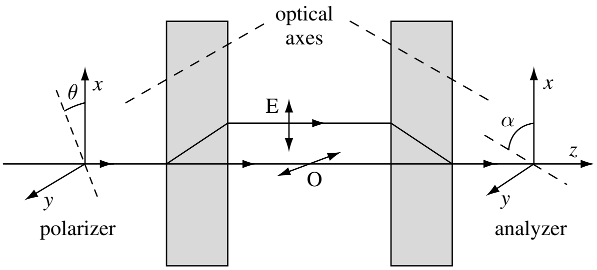

Fig. 3.3. Decomposition and recombination of polarizations using birefringent plates.

originating from the difference between the ordinary and extraordinary indices in the birefringent plates (equivalently, we can imagine that this difference is cancelled by an intermediate birefringent plate which is oriented appropriately; see Exercise 3.3.1). Under these conditions the light wave at the exit of the second birefringent plate is polarized in the $\hat{n}_\theta$ direction. The recombination of the two x and y beams gives the initial light beam polarized in the $\hat{n}_\theta$ direction, and the intensity at the exit of the analyzer is reduced as before by a factor $\cos^2(\theta-\alpha)$.

If we limit ourselves to linear polarization states, we can describe any polarization state as a real unit vector in the xOy plane, in which a possible orthonormal basis is constructed from the vectors $|x\rangle$ and $|y\rangle$. However, if we want to describe an arbitrary polarization, we need to introduce a two-dimensional *complex* vector space $\mathcal{H}$. This space will be the *vector space of the polarization states*. Let us return to the general case (3.1), introducing complex notation $\vec{\mathcal{E}} = (\mathcal{E}_x,\ \mathcal{E}_y)$ for the wave amplitudes:

$$\mathcal{E}_x = E_{0x}e^{i\delta_x}, \quad \mathcal{E}_y = E_{0y}e^{i\delta_y}, \tag{3.6}$$

which allows us to write (3.1) in the form

$$\begin{aligned} E_x(t) &= E_{0x}\cos(\omega t - \delta_x) = \mathrm{Re}\left(E_{0x}e^{i\delta_x}e^{-i\omega t}\right) = \mathrm{Re}\left(\mathcal{E}_x e^{-i\omega t}\right), \\ E_y(t) &= E_{0y}\cos(\omega t - \delta_y) = \mathrm{Re}\left(E_{0y}e^{i\delta_y}e^{-i\omega t}\right) = \mathrm{Re}\left(\mathcal{E}_y e^{-i\omega t}\right). \end{aligned} \tag{3.7}$$

We have already noted that owing to the arbitrariness of the time origin, only the relative phase $\delta = (\delta_y - \delta_x)$ is physically relevant and we can multiply $\mathcal{E}_x$ and $\mathcal{E}_y$ by a common phase factor $\exp(i\beta)$ without any physical consequences. For example, it is always possible to choose $\delta_x = 0$. The light intensity is given by (3.2):

$$\mathcal{I} = k(|\mathcal{E}_x|^2 + |\mathcal{E}_y|^2) = k|\vec{\mathcal{E}}|^2 = kE_0^2. \tag{3.8}$$

An important special case of (3.7) is that of *circular polarization*, where $E_{0x} = E_{0y} = E_0/\sqrt{2}$ and $\delta_y = \pm\pi/2$ (we have conventionally chosen $\delta_x = 0$). If $\delta_y = +\pi/2$, the tip of the electric field vector traces a circle in the xOy plane in the counterclockwise sense. The components $E_x(t)$ and $E_y(t)$ are given by

$$\begin{aligned} E_x(t) &= \mathrm{Re}\left(\frac{E_0}{\sqrt{2}}\mathrm{e}^{-\mathrm{i}\omega t}\right) = \frac{E_0}{\sqrt{2}}\cos\omega t, \\ E_y(t) &= \mathrm{Re}\left(\frac{E_0}{\sqrt{2}}\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{e}^{\mathrm{i}\pi/2}\right) = \frac{E_0}{\sqrt{2}}\cos(\omega t - \pi/2) = \frac{E_0}{\sqrt{2}}\sin\omega t. \end{aligned} \tag{3.9}$$

An observer at whom the light wave arrives sees the tip of the electric field vector tracing a circle of radius $E_0/\sqrt{2}$ counterclockwise in the xOy plane. The corresponding polarization is termed *right-handed circular polarization*.[2] When $\delta_y = -\pi/2$, we obtain *left-handed circular polarization* – the circle is traced in the clockwise sense:

$$\begin{aligned} E_x(t) &= \mathrm{Re}\left(\frac{E_0}{\sqrt{2}}\mathrm{e}^{-\mathrm{i}\omega t}\right) = \frac{E_0}{\sqrt{2}}\cos\omega t, \\ E_y(t) &= \mathrm{Re}\left(\frac{E_0}{\sqrt{2}}\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{e}^{-\mathrm{i}\pi/2}\right) = \frac{E_0}{\sqrt{2}}\cos(\omega t + \pi/2) = -\frac{E_0}{\sqrt{2}}\sin\omega t. \end{aligned} \tag{3.10}$$

These right- and left-handed circular polarization states are obtained experimentally starting from linear polarization at an angle of 45° to the axes and then introducing a phase shift $\pm\pi/2$ of the field in the Ox or Oy direction by means of a quarter-wave plate.

In complex notation the fields $\mathcal{E}_x$ and $\mathcal{E}_y$ are written as

$$\mathcal{E}_x = \frac{1}{\sqrt{2}}E_0, \quad \mathcal{E}_y = \frac{1}{\sqrt{2}}E_0\mathrm{e}^{\pm\mathrm{i}\pi/2} = \frac{\pm\mathrm{i}}{\sqrt{2}}E_0,$$

where the $+$ sign corresponds to right-handed circular polarization and the $-$ to left-handed. The proportionality factor E_0 common to $\mathcal{E}_x$ and $\mathcal{E}_y$ defines the intensity of the light wave and plays no role in describing the polarization, which is characterized by the normalized vectors

$$\boxed{|\mathrm{R}\rangle = -\frac{1}{\sqrt{2}}(|x\rangle + \mathrm{i}|y\rangle), \quad |\mathrm{L}\rangle = \frac{1}{\sqrt{2}}(|x\rangle - \mathrm{i}|y\rangle)}\,. \tag{3.11}$$

The overall minus sign in the definition of $|\mathrm{R}\rangle$ has been introduced to be consistent with the conventions of Chapter 10. Equation (3.11) shows that the mathematical description of polarization leads naturally to the use of unit vectors in a complex two-dimensional vector space $\mathcal{H}$, in which the vectors $|x\rangle$ and $|y\rangle$ form one possible orthonormal basis.

[2] See Fig. 10.8. Our definition of right- and left-handed circular polarization is the one used in elementary particle physics. With this definition, right- (left-) handed circular polarization corresponds to positive (negative) helicity, that is, to projection of the photon spin on the direction of propagation equal to $+\hbar$ ($-\hbar$). However, this definition is not universal; optical physicists often use the opposite, but, as one of them has remarked (E. Hecht, *Optics*, New York: Addison-Wesley (1987), Chapter 8): "This choice of terminology is admittedly a bit awkward. Yet its use in optics is fairly common, even though it is completely antithetic to the more reasonable convention adopted in elementary particle physics."

Above we have established the correspondence between linear polarization in the $\hat{n}_\theta$ direction and the unit vector $|\theta\rangle$ of $\mathcal{H}$, as well as the correspondence between the two circular polarizations and the two vectors (3.11) of $\mathcal{H}$. We are now going to generalize this correspondence by constructing the polarization corresponding to the most general normalized vector $|\Phi\rangle$ of $\mathcal{H}$:[3]

$$|\Phi\rangle = \lambda|x\rangle + \mu|y\rangle, \quad |\lambda|^2 + |\mu|^2 = 1. \tag{3.12}$$

It is always possible to choose λ to be real (in Exercise 3.3.2 we show that the physics is unaffected if λ is complex). The numbers λ and μ can then be parametrized by two angles θ and η:

$$\lambda = \cos\theta, \quad \mu = \sin\theta e^{i\eta}.$$

We shall imagine a device containing two birefringent plates and a linear polarizer, on which an electromagnetic wave (3.7) is incident. This device will be called a (λ, μ) *polarizer*.

- The first birefringent plate changes the phase of $\mathcal{E}_y$ by $-\eta$ while leaving $\mathcal{E}_x$ unchanged:

$$\mathcal{E}_x \to \mathcal{E}_x^{(1)} = \mathcal{E}_x, \quad \mathcal{E}_y \to \mathcal{E}_y^{(1)} = \mathcal{E}_y e^{-i\eta}.$$

- The linear polarizer projects on the $\hat{n}_\theta$ direction:

$$\begin{aligned}\vec{\mathcal{E}}^{(1)} \to \vec{\mathcal{E}}^{(2)} &= \left(\mathcal{E}_x^{(1)} \cos\theta + \mathcal{E}_y^{(1)} \sin\theta\right)\hat{n}_\theta \\ &= \left(\mathcal{E}_x \cos\theta + \mathcal{E}_y \sin\theta e^{-i\eta}\right)\hat{n}_\theta.\end{aligned}$$

- The second birefringent plate leaves $\mathcal{E}_x^{(2)}$ unchanged and shifts the phase of $\mathcal{E}_y^{(2)}$ by η:

$$\mathcal{E}_x^{(2)} \to \mathcal{E}'_x = \mathcal{E}_x^{(2)}, \quad \mathcal{E}_y^{(2)} \to \mathcal{E}'_y = \mathcal{E}_y^{(2)} e^{i\eta}.$$

The combination of the three operations is represented by the transformation $\vec{\mathcal{E}} \to \vec{\mathcal{E}}'$ which can be written in terms of components:

$$\begin{aligned}\mathcal{E}'_x &= \mathcal{E}_x \cos^2\theta + \mathcal{E}_y \sin\theta\cos\theta e^{-i\eta} = |\lambda|^2\mathcal{E}_x + \lambda\mu^*\mathcal{E}_y, \\ \mathcal{E}'_y &= \mathcal{E}_x \sin\theta\cos\theta e^{i\eta} + \mathcal{E}_y \sin^2\theta = \lambda^*\mu\mathcal{E}_x + |\mu|^2\mathcal{E}_y.\end{aligned} \tag{3.13}$$

The operation (3.13) amounts to *projection* on $|\Phi\rangle$. In fact, if we choose to write the vectors $|x\rangle$ and $|y\rangle$ as column vectors

$$|x\rangle = \begin{pmatrix}1\\0\end{pmatrix}, \quad |y\rangle = \begin{pmatrix}0\\1\end{pmatrix}, \tag{3.14}$$

then the projector $\mathcal{P}_\Phi$

$$\mathcal{P}_\Phi = |\Phi\rangle\langle\Phi| = \big(\lambda|x\rangle + \mu|y\rangle\big)\big(\lambda^*\langle x| + \mu^*\langle y|\big)$$

[3] We shall use upper-case letters $|\Phi\rangle$ or $|\Psi\rangle$ for generic vectors of $\mathcal{H}$ of the form (3.12) or (3.16), to avoid any confusion with an angle, as for $|\theta\rangle$ or $|\alpha\rangle$.

is represented by the matrix

$$\mathcal{P}_\Phi = \begin{pmatrix} |\lambda|^2 & \lambda\mu^* \\ \lambda^*\mu & |\mu|^2 \end{pmatrix}. \tag{3.15}$$

We can put the incident field $\vec{\mathcal{E}}$ (3.7) in correspondence with a (non-normalized) vector $|\mathcal{E}\rangle$ of $\mathcal{H}$ with the complex components $\mathcal{E}_x$ and $\mathcal{E}_y$:

$$|\mathcal{E}\rangle = \mathcal{E}_x|x\rangle + \mathcal{E}_y|y\rangle.$$

Using $|\mathcal{E}\rangle$ we can define a vector $|\Psi\rangle$ normalized to unity by $|\mathcal{E}\rangle = E_0|\Psi\rangle$:

$$|\Psi\rangle = \nu|x\rangle + \sigma|y\rangle, \quad |\nu|^2 + |\sigma|^2 = 1, \tag{3.16}$$

where

$$\nu = \frac{\mathcal{E}_x}{E_0}, \quad \sigma = \frac{\mathcal{E}_y}{E_0}.$$

The normalized vector $|\Psi\rangle$ which describes the polarization of the wave (3.7) is called the *Jones vector*. According to (3.13) and (3.15), the electric field at the exit of the (λ, μ) polarizer will be

$$|\mathcal{E}'\rangle = \mathcal{P}_\Phi|\mathcal{E}\rangle = E_0\mathcal{P}_\Phi|\Psi\rangle = E_0|\Phi\rangle\langle\Phi|\Psi\rangle. \tag{3.17}$$

Now let us generalize everything we have obtained for the linear polarizer to the (λ, μ) polarizer. The latter projects the polarization state $|\Psi\rangle$ onto $|\Phi\rangle$ with amplitude equal to $\langle\Phi|\Psi\rangle$:

$$\overline{a}(\Psi \to \Phi) = \langle\Phi|\Psi\rangle. \tag{3.18}$$

At the exit of the polarizer the intensity is reduced by a factor $|\overline{a}(\Psi \to \Phi)|^2 = |\langle\Phi|\Psi\rangle|^2$. If the polarization state is described by the unit vector $|\Phi\rangle$ (3.12), then the transmission through the (λ, μ) polarizer is 100%. On the other hand, the polarization state

$$|\Phi_\perp\rangle = -\mu^*|x\rangle + \lambda^*|y\rangle \tag{3.19}$$

is completely stopped by the (λ, μ) polarizer. The polarization state (3.16) is in general an elliptic polarization. It is easy to determine the characteristics of the corresponding ellipse and the direction in which it is traced (Exercise 3.3.2).

The states $|\Phi\rangle$ and $|\Phi_\perp\rangle$ form an orthonormal basis of $\mathcal{H}$ obtained from the $\{|x\rangle, |y\rangle\}$ basis by a unitary transformation U:

$$U = \begin{pmatrix} \lambda & \mu \\ -\mu^* & \lambda^* \end{pmatrix}.$$

In summary, we have shown that any polarization state can be put into correspondence with a normalized vector $|\Phi\rangle$ of a two-dimensional complex space $\mathcal{H}$. The vectors $|\Phi\rangle$ and $\exp(\mathrm{i}\beta)|\Phi\rangle$ represent the same polarization state. Stated more precisely, a polarization state can be put into correspondence with a vector up to a phase.

3.1.2 The photon polarization

Now we shall show that the mathematical formalism used above to describe the polarization of a light wave can be carried over without modification to the description of the polarization of a photon. However, the fact that the mathematical formalism is identical in the two cases should not obscure the fact that the physical interpretation is radically modified. We shall return to the experiment of Fig. 3.2 and reduce the light intensity such that individual photons are registered by the photomultipliers D_x and D_y, which respectively detect photons polarized in the Ox and Oy directions. We then observe the following:

- only one of the two photomultipliers is triggered by a photon incident on the plate. Like the neutrons of Chapter 1, the photons arrive in lumps: they are never split.
- the probability p_x (p_y) of D_x (D_y) being triggered by a photon incident on the plate is $\mathsf{p}_x = \cos^2\theta$ ($\mathsf{p}_y = \sin^2\theta$).

This result must hold true if we want to recover classical optics in the limit where the number N of photons is large. In fact, if N_x and N_y are the numbers of photons detected by D_x and D_y, we must have

$$\mathsf{p}_x = \lim_{N\to\infty} \frac{N_x}{N}, \quad \mathsf{p}_y = \lim_{N\to\infty} \frac{N_y}{N}$$

and $\mathcal{I}_x \propto N_x = N\cos^2\theta$, $\mathcal{I}_y \propto N_y = N\sin^2\theta$ in the limit $N \to \infty$. However, the fate of an individual photon cannot be predicted. We can only know its *probability* of detection by D_x or D_y. The need to resort to probabilities is an intrinsic feature of quantum physics, whereas in classical physics resorting to probabilities is only a way to take into account the complexity of a phenomenon whose details we cannot (or do not want to) know. For example, when flipping a coin, complete knowledge of the initial conditions under which the coin is thrown and inclusion of the air resistance, the state of the ground on which the coin lands, etc. permit us in principle to predict the result. Some physicists[4] have suggested that the probabilistic nature of quantum mechanics has an analogous origin: if we had access to additional variables which at present we do not know, the so-called *hidden variables*, we would be able to predict with certainty the fate of each individual photon. This hidden variable hypothesis has some utility in discussions of the foundations of quantum physics. Nevertheless, in Chapter 6 we shall see that, given very plausible hypotheses, such variables are excluded by experiment.

However, probabilities alone provide only a very incomplete description of the photon polarization. A complete description requires also the introduction of *probability amplitudes*. Probability amplitudes, which we denote a (the difference between the wave amplitudes of the preceding subsection and probability amplitudes is emphasized by using different notation: a instead of $\overline{a}$), are complex numbers, and probabilities correspond to their squared modulus $|a|^2$. To make manifest the incomplete nature of probabilities

[4] Including de Broglie and Bohm.

alone, let us again consider the apparatus of Fig. 3.3. Between the two plates a photon follows either the trajectory of an extraordinary ray polarized in the Ox direction, called an x trajectory, or the trajectory of an ordinary ray polarized in the Oy direction, called a y trajectory. According to purely probabilistic reasoning, a photon following an x trajectory has probability $\cos^2\theta\cos^2\alpha$ of being transmitted by the analyzer, and a photon following a y trajectory has the corresponding probability $\sin^2\theta\sin^2\alpha$. The total probability for a photon to be transmitted by the analyzer is therefore

$$\mathsf{p}_{\text{tot}} = \cos^2\theta\cos^2\alpha + \sin^2\theta\sin^2\alpha. \tag{3.20}$$

This is not what is found from experiment, which confirms the result obtained earlier using wave arguments:

$$\mathsf{p}_{\text{tot}} = \cos^2(\theta-\alpha).$$

A correct reasoning must be based on probability amplitudes, just as before we used wave amplitudes. Probability amplitudes obey the same rules as wave amplitudes, which guarantees that the results of optics are reproduced when the number of photons $N\to\infty$. The probability amplitude for a photon linearly polarized in the $\hat{n}_\theta$ direction to be polarized in the $\hat{n}_\alpha$ direction is given by (3.4): $a(\theta\to\alpha) = \cos(\theta-\alpha) = \hat{n}_\theta\cdot\hat{n}_\alpha$. We obtain the following table of probability amplitudes for the experiment of Fig. 3.3:

$$a(\theta\to x) = \cos\theta, \quad a(x\to\alpha) = \cos\alpha,$$

$$a(\theta\to y) = \sin\theta, \quad a(y\to\alpha) = \sin\alpha.$$

This example provides an illustration of the rules governing the combination of probability amplitudes. The probability amplitude a_x for an incident photon following an x trajectory to be transmitted by the analyzer is

$$a_x = a(\theta\to x)a(x\to\alpha) = \cos\theta\cos\alpha.$$

This expression suggests the *factorization* rule for amplitudes: a_x is the product of the amplitudes $a(\theta\to x)$ and $a(x\to\alpha)$. This factorization rule guarantees that the corresponding rule for the probabilities holds. We also have

$$a_y = a(\theta\to y)a(y\to\alpha) = \sin\theta\sin\alpha.$$

If the experimental setup does not allow us to know which trajectory a photon has followed, the amplitudes must be added. The total probability amplitude for a photon to be transmitted by the analyzer is then

$$a_{\text{tot}} = a_x + a_y = \cos\theta\cos\alpha + \sin\theta\sin\alpha = \cos(\theta-\alpha), \tag{3.21}$$

and the corresponding probability is $\cos^2(\theta-\alpha)$, in agreement with the result (3.5) of classical optics. If there is a way to distinguish between the two trajectories, the interference is destroyed and the probabilities must be added as in (3.20).

Since the rules for combining probability amplitudes are the same as those for wave amplitudes, these rules will apply if the polarization state of a photon is described by a normalized vector in a two-dimensional vector space $\mathcal{H}$, called the *space of states*. In the present case this is the space of polarization states. When a photon is linearly polarized in the Ox (Oy) direction, we can put this polarization state in correspondence with a vector $|x\rangle$ ($|y\rangle$) of this space. Such a polarization state is obtained by allowing a photon to pass through a linear polarizer oriented in the Ox (Oy) direction. The probability that a photon polarized in the Ox direction will be transmitted by an analyzer oriented in the Oy direction is zero: the probability amplitude $a(x \to y) = 0$. Conversely, the probability that a photon polarized in the Ox or Oy direction will be transmitted by an analyzer oriented in the same direction is equal to unity, and so

$$|a(x \to x)| = |a(y \to y)| = 1, \quad a(x \to y) = a(y \to x) = 0.$$

These relations are satisfied if $|x\rangle$ and $|y\rangle$ form an orthonormal basis of $\mathcal{H}$ and if we identify the probability amplitudes as scalar products:

$$a(x \to x) = \langle x|x\rangle = 1, \quad a(y \to y) = \langle y|y\rangle = 1, \quad a(y \to x) = \langle x|y\rangle = 0. \tag{3.22}$$

The most general *linear* polarization state is the state in which the polarization makes an angle θ with Ox. This state will be represented by the vector

$$|\theta\rangle = \cos\theta|x\rangle + \sin\theta|y\rangle. \tag{3.23}$$

Equations (3.22) and (3.23) ensure that the probability amplitudes listed above are correctly given by the scalar products, for example,

$$a(\theta \to x) = \langle x|\theta\rangle = \cos\theta,$$

or, in general, if $|\alpha\rangle$ is a state of linear polarization,

$$a(\theta \to \alpha) = \langle \alpha|\theta\rangle = \cos(\theta - \alpha).$$

The most general polarization state will be described by a normalized vector called a *state vector*:

$$|\Phi\rangle = \lambda|x\rangle + \mu|y\rangle, \quad |\lambda|^2 + |\mu|^2 = 1.$$

As in the wave case, the vectors $|\Phi\rangle$ and $\exp(i\beta)|\Phi\rangle$ represent the same physical state: a physical state is represented by a vector up to a phase in the space of states. The *probability amplitude for finding a polarization state* $|\Psi\rangle$ *in* $|\Phi\rangle$ will be given by the scalar product $\langle \Phi|\Psi\rangle$, and the projection onto a given polarization state will be realized by the (λ, μ) polarizer described in the preceding subsection. In summary, we have used a specific example, that of the polarization of a photon, to illustrate the construction of the Hilbert space of states.

The photon polarization along some (complex) direction is an example of a *quantum physical property*. The interpretation of a quantum physical property differs radically from that of a classical physical property. We shall illustrate this by examining the photon polarization. At first we limit ourselves to the simplest case, that of a linear polarization state. Using a linear polarizer oriented in the Ox direction, we prepare an ensemble of photons all in the state $|x\rangle$. The photons arrive one by one at the polarizer, and all the photons which are transmitted by the polarizer are in the state $|x\rangle$. This is the stage of *preparation of the quantum system*, where one only keeps the photons which have passed through the polarizer aligned in the Ox direction. The next stage, the *test* stage, consists of testing this polarization by allowing the photons to pass through a linear analyzer. If the analyzer is parallel to Ox the photons are transmitted with unit probability and if it is parallel to Oy they are transmitted with zero probability. In both cases the result of the test can be predicted with certainty. The physical property "polarization of a photon prepared in the state $|x\rangle$" takes well-defined values if the basis $\{|x\rangle, |y\rangle\}$ is chosen for the test. On the other hand, if we use analyzers oriented in the direction $\hat{n}_\theta$ corresponding to the state $|\theta\rangle$ (3.23) and in the perpendicular direction $\hat{n}_{\theta_\perp}$ corresponding to the state

$$|\theta_\perp\rangle = -\sin\theta|x\rangle + \cos\theta|y\rangle, \tag{3.24}$$

we can predict only the transmission probability $|\langle\theta|x\rangle|^2 = \cos^2\theta$ in the first case and $|\langle\theta_\perp|x\rangle|^2 = \sin^2\theta$ in the second. The physical property "polarization of the photon in the state $|x\rangle$" has no well-defined value in the basis $\{|\theta\rangle, |\theta_\perp\rangle\}$. In other words, the physical property "polarization" is associated with a given basis, and the two bases $\{|x\rangle, |y\rangle\}$ and $\{|\theta\rangle, |\theta_\perp\rangle\}$ are termed *incompatible* (except when $\theta = 0$ and $\theta = \pi/2$). *Complementary bases* are a special case of incompatible ones: in a Hilbert space of dimension N, two bases $\{|m\rangle\}$ and $\{|\mu\rangle\}$ are termed complementary if $|\langle m|\mu\rangle|^2 = 1/N$ for all m and μ.

The preceding discussion should be made more precise in two respects. First, it is clearly impossible to test the polarization of an isolated photon. The polarization test requires that we are provided with a number $N \gg 1$ of photons prepared under identical conditions. Let us then suppose that N photons have been prepared in a certain polarization state and that they are tested by a linear analyzer oriented in the Ox direction. If we find – within the experimental accuracy of the apparatus – that the photons pass through the analyzer with a probability of 100%, we can deduce that the photons have been prepared in the state $|x\rangle$. The observation of a single photon obviously does not allow us to arrive at this conclusion, unless we know beforehand in which basis it was prepared. The second point is that even if the photons are transmitted with a probability $\cos^2\theta$, we cannot deduce that they have been prepared in the linear polarization state (3.23). In fact, we will observe the same transmission probability if the photons have been prepared in an elliptic polarization state (3.12) with

$$\lambda = \cos\theta e^{i\delta_x}, \quad \mu = \sin\theta e^{i\delta_y}.$$

Only a test whose results have probability 0 or 1 allows the photon polarization state to be determined unambiguously with one orientation of the analyzer. Otherwise, a second orientation will be necessary to determine the phases.

In the representation (3.14) of the basis vectors of $\mathcal{H}$, the projectors $\mathcal{P}_x$ and $\mathcal{P}_y$ onto the states $|x\rangle$ and $|y\rangle$ are represented by matrices

$$\mathcal{P}_x = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathcal{P}_y = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

which commute: $[\mathcal{P}_x, \mathcal{P}_y] = 0$. The two operators are compatible according to the definition of Section 2.3.3. The projectors $\mathcal{P}_\theta$ and $\mathcal{P}_{\theta_\perp}$ can be calculated directly from (3.15):

$$\mathcal{P}_\theta = \begin{pmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{pmatrix}, \quad \mathcal{P}_{\theta_\perp} = \begin{pmatrix} \sin^2\theta & -\sin\theta\cos\theta \\ -\sin\theta\cos\theta & \cos^2\theta \end{pmatrix}.$$

They commute with each other, but not with either $\mathcal{P}_x$ or $\mathcal{P}_y$: $\mathcal{P}_x$ and $\mathcal{P}_\theta$, for example, are incompatible. The commutation (or noncommutation) of operators is the mathematical translation of the compatibility (or incompatibility) of physical properties.

As another choice of basis we can use the right- and left-handed circular polarization states $|\mathrm{R}\rangle$ and $|\mathrm{L}\rangle$ of (3.11). The basis $\{|\mathrm{R}\rangle, |\mathrm{L}\rangle\}$ is incompatible with any basis constructed using linear polarization states, and in fact complementary to any such basis. The projectors $\mathcal{P}_\mathrm{R}$ and $\mathcal{P}_\mathrm{L}$ onto these circular polarization states are

$$\mathcal{P}_\mathrm{R} = \frac{1}{2}\begin{pmatrix} 1 & -\mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix}, \quad \mathcal{P}_\mathrm{L} = \frac{1}{2}\begin{pmatrix} 1 & \mathrm{i} \\ -\mathrm{i} & 1 \end{pmatrix}. \tag{3.25}$$

We can use $\mathcal{P}_\mathrm{R}$ and $\mathcal{P}_\mathrm{L}$ to construct the remarkable Hermitian operator Σ_z:

$$\boxed{\Sigma_z = \mathcal{P}_\mathrm{R} - \mathcal{P}_\mathrm{L} = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}}\,. \tag{3.26}$$

This operator has the states $|\mathrm{R}\rangle$ and $|\mathrm{L}\rangle$ as its eigenvectors, and their respective eigenvalues are $+1$ and -1:

$$\Sigma_z|\mathrm{R}\rangle = |\mathrm{R}\rangle, \quad \Sigma_z|\mathrm{L}\rangle = -|\mathrm{L}\rangle. \tag{3.27}$$

This result suggests that the Hermitian operator Σ_z with eigenvectors $|\mathrm{R}\rangle$ and $|\mathrm{L}\rangle$ is associated with the physical property called "circular polarization." We shall see in Chapter 10 that $\hbar\Sigma_z = J_z$ is the operator representing the physical property called "z component of the photon angular momentum (or spin)." We also observe that $\exp(-\mathrm{i}\theta\Sigma_z)$ is an operator which performs rotations by an angle θ about the Oz axis, as can be seen from a simple calculation (Exercise 3.3.3)

$$\exp(-\mathrm{i}\theta\Sigma_z) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{3.28}$$

and $\exp(-\mathrm{i}\theta\Sigma_z)$ transforms the state $|x\rangle$ into the state $|\theta\rangle$ and $|y\rangle$ into $|\theta_\perp\rangle$:

$$\exp(-\mathrm{i}\theta\Sigma_z)|x\rangle = |\theta\rangle, \quad \exp(-\mathrm{i}\theta\Sigma_z)|y\rangle = |\theta_\perp\rangle. \tag{3.29}$$

3.1.3 Quantum cryptography

Quantum cryptography is a recent invention based on the incompatibility of two different bases of linear polarization states. Ordinary cryptography makes use of an encryption key known only to the transmitter and receiver. This is called *secret-key* cryptography. It is in principle very secure,[5] but it is necessary that the transmitter and receiver be able to exchange the key without its being intercepted by a spy. The key must be changed often, because a set of messages encoded using the same key can reveal regularities which permit decipherment by a third party. The process of transmitting a secret key is risky, and for this reason it is preferable to use systems based on a different principle, the so-called *public-key* systems, where the key is made public, for example via the Internet. A public-key system currently in use is based on the difficulty of factoring a very large number N into primes,[6] whereas the reverse operation is straightforward: without a calculator one can obtain $137 \times 53 = 7261$ in a few seconds, but given 7261 it would take some time to factor it into primes. The number of instructions needed for a computer using the best modern algorithms to factor a number N into primes grows with N roughly as $\exp[(\ln N)^{1/3}]$.[7] In a public-key system, the receiver, conventionally named Bob, publicly sends to the transmitter, conventionally named Alice, a very large number $N = pq$ which is the product of two primes p and q, as well a number c having no common factor with $(p-1)(q-1)$. Knowledge of N and c is sufficient for Alice to encrypt the message, but decipherment requires knowing the numbers p and q. Of course, a spy, conventionally named Eve, possessing a sufficiently powerful computer and enough time can manage to crack the code, but in general one can count on keeping the contents of the message secret for a limited period of time. However, it is not impossible that eventually very powerful algorithms will be found for factoring a number into primes, and, moreover, if quantum computers (Section 6.4.2) ever see the light of day, they will push the limits of factorization very far. Fortunately, thanks to quantum mechanics we are nearly at the point of being able to counteract the efforts of spies.

"Quantum cryptography" is a catchy phrase, but somewhat inaccurate. The point is not that a message is encrypted using quantum physics, but rather that quantum physics is used to ensure that the key has been transmitted securely: a more accurate terminology is thus "quantum key distribution" (QKD). A message, encrypted or not, can be transmitted using the two orthogonal linear polarization states of a photon, for example, $|x\rangle$ and $|y\rangle$. We can adopt the convention of assigning the value 1 to the polarization $|x\rangle$ and 0 to the polarization $|y\rangle$; then each photon transports a bit of information. The entire message, encrypted or not, can be written in binary code, that is, as a series of ones and zeros, and the message 1001110 can be encoded by Alice using the photon sequence $xyyxxxy$ and then sent to Bob via, for example, an optical fiber. Using a birefringent plate, Bob

[5] An absolutely secure encryption was discovered by Vernam in 1917. However, absolute security requires that the key be as long as the message and that it be used only a single time!

[6] Called RSA encryption, discovered by Rivest, Shamir, and Adleman in 1977.

[7] At present the best factorization algorithm requires a number of operations $\sim \exp[1.9(\ln N)^{1/3}(\ln\ln N)^{2/3}]$. One cannot hope to factor numbers with more than 180 figures ($\sim 10^{20}$ instructions) in a reasonable amount of time.

will separate the photons of vertical and horizontal polarization as in Fig. 3.2, and two detectors located behind the plate will permit him to decide if a photon was horizontally or vertically polarized. In this way he can reconstruct the message. If this were an ordinary message, there would of course be much simpler and more efficient methods of sending it! At this point, let us just note that if Eve eavesdrops on the fiber, detects the photons and their polarization, and then sends to Bob other photons with the same polarization as the ones sent by Alice, Bob is none the wiser. The situation would be the same for any device functioning in a classical manner, that is, any device that does not use the superposition principle: if the spy takes sufficient precautions, the spying is undetectable, because she can send a signal that is arbitrarily close to the original one.

This is where quantum mechanics and the superposition principle come to the aid of Alice and Bob, allowing them to be sure that their message has not been intercepted. The message need not be long (the method of transmission via polarization is not very efficient). The idea in general is to transmit the key permiting encryption of a later message, a key which can be replaced when necessary. Alice sends Bob four types of photon: photons polarized along Ox ($\updownarrow$) and Oy ($\leftrightarrow$) as before, and photons polarized along axes rotated by $\pm 45^\circ$, that is, Ox' (⤡) and Oy' (⤢), respectively corresponding to bits 1 and 0. Again Bob analyzes the photons sent by Alice, now using analyzers oriented in four directions, vertical/horizontal and $\pm 45^\circ$. One possibility is to use a birefringent crystal randomly oriented vertically or at 45° from the vertical and to detect the photons leaving this crystal as in Fig. 3.3. However, instead of rotating the crystal+detector ensemble, it is easier to use a Pockels cell, which allows a given polarization to be transformed into one of arbitrary orientation while keeping the crystal+detector ensemble fixed (Fig. 3.4). Bob records 1 if the photon has polarization $\updownarrow$ or ⤡, and 0 if it has polarization $\leftrightarrow$ or ⤢. After recording a sufficient number of photons, Bob announces publicly the analyzer sequence he has used, but not his results. Alice compares her polarizer sequence to that of Bob and also publicly gives him the list of polarizers compatible with his analyzers. The bits corresponding to incompatible analyzers and polarizers are rejected ($-$), and, for the other bits, Alice and Bob are certain that their values are the same. It is these bits which will serve to construct the key, and they are known only to Bob and Alice, because an outsider knows only the list of orientations and not the results. An example of photon exchanges between Alice and Bob is given in Fig. 3.5.

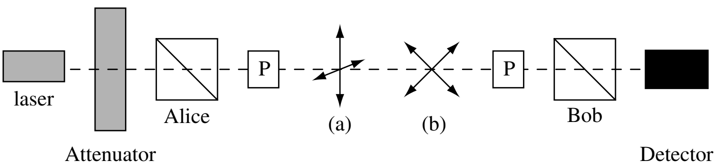

Fig. 3.4. The BB84 protocol. An attenuted laser beam allows Alice to send individual photons. A birefringent crystal selects a given linear polarization, which can be rotated thanks to a Pockels cell P. The photons are polarized, either vertically/horizontally (a), or to $\pm 45^\circ$ (b).

Alice's polarizers	↕	↔	⤢	↕	⤢	⤢	⤡	↕	⤡
sequence of bits	1	0	0	1	0	0	1	1	1
Bob's analyzers	+	×	+	+	×	×	+	+	×
Bob's measurements	1	1	0	1	0	0	1	1	1
retained bits	1	–	–	1	0	0	–	1	1

Fig. 3.5. Quantum cryptography: transmission of polarized photons between Bob and Alice.

The only thing left is to ensure that the message has not been intercepted and that the key it contains can be used without risk. Alice and Bob randomly choose a subset of their key and compare it publicly. If Eve has intercepted the photons, this will result in a reduction of the correlation between the values of their bits. Suppose, for example, that Alice sends a photon polarized in the Ox direction. If Eve intercepts it using a polarizer oriented in the Ox' direction, and if the photon is transmitted by her analyzer, she does not know that this photon was initially polarized along the Ox direction, and so she resends Bob a photon polarized in the Ox' direction, and in 50% of cases Bob will not obtain the right result. Since Eve has one chance in two of orienting her analyzer in the right direction, Alice and Bob will register a difference in 25% of cases and conclude that the message has been intercepted. The use of two complementary bases maximizes the security of the BB84 protocol. Of course, this discussion is greatly simplified. It does not take into account the possibilities of errors which must be corrected, and moreover it is based on recording impacts of isolated photons, while in practice one sends packets of coherent states with a small ($\langle n \rangle \sim 0.1$) average number of photons by using an attenuated laser beam.[8] Nevertheless, the method is correct in principle, and, to this day, two devices capable of realizing transmissions over several tens of kilometers are available on the market.

3.2 Spin 1/2

3.2.1 Angular momentum and magnetic moment in classical physics

Our second example of an elementary quantum system will be that of spin 1/2. Since for such a system there is no classical wave limit as there is in the case of the photon, our classical discussion will be much shorter than that of the preceding section. We consider a particle of mass m and charge q describing a closed orbit in the field of a central force (Fig. 3.6). We denote the position and momentum of this particle as $\vec{r}(t)$ and $\vec{p}(t)$.

[8] In the case of the transmission of isolated photons, the theorem of quantum cloning (Section 6.4.2) guarantees that it is impossible for Eve to fool Bob. However, Eve can slightly reduce her error rate by using a more sophisticated method: see Exercise 15.5.3.