2

What is a qubit?

2.1 The polarization of light

Our first example of a qubit will be the polarization of a photon. First we briefly review the subject of light polarization. The *polarization of light* was demonstrated for the first time by the Chevalier Malus in 1809. He observed the light of the setting sun reflected by the glass of a window in the Luxembourg Palace in Paris through a crystal of Iceland spar. He showed that when the crystal was rotated, one of the two images of the sun disappeared. Iceland spar is a birefringent crystal which, as we shall see below, decomposes a light ray into two rays polarized in perpendicular directions, while the ray reflected from the glass is (partially) polarized. When the crystal is suitably oriented one then observes the disappearance (or strong attenuation) of one of the two rays. The phenomenon of polarization displays the vector nature of light waves, a property which is shared by shear sound waves: in an isotropic crystal, a sound wave can correspond either to a vibration transverse to the direction of propagation, i.e., a shear wave, or to a longitudinal vibration, i.e., a compression wave. In the case of light the vibration is only transverse: the electric field of a light wave is orthogonal to the propagation direction.

Let us recall the mathematical description of a planar and monochromatic scalar wave traveling in the z direction. The amplitude of vibration $u(z, t)$ as a function of time t has the form

$$u(z, t) = u_0 \cos(\omega t - kz),$$

where ω is the vibrational frequency, k is the wave vector ($k = 2\pi/\lambda$, where λ is the wavelength), related by $\omega = ck$, where c is the propagation speed, here the speed of light. It can be immediately checked that a maximum of $u(z, t)$ moves at speed $\omega/k = c$. In what follows we shall always work in a plane at fixed z, for example, the $z = 0$ plane where

$$u(z = 0, t) := u(t) = u_0 \cos \omega t.$$

https://doi.org/10.1017/CBO9780511755361.004

When an electromagnetic wave passes through a polarizing filter (a *polarizer*), the vibration transmitted by the filter is a vector in the xOy plane transverse to the propagation direction:

$$\begin{aligned} E_x &= E_0 \cos\theta \cos\omega t, \\ E_y &= E_0 \sin\theta \cos\omega t, \end{aligned} \tag{2.1}$$

where θ depends on the orientation of the filter. The light intensity (or energy) measured, for example, using a photoelectric cell is proportional to the squared electric field $I \propto E_0^2$ (in general, the energy of a vibration is proportional to the squared vibrational amplitude). The unit vector [1] $\hat{p}$ in the xOy plane

$$\hat{p} = (\cos\theta, \sin\theta), \quad \vec{E} = E_0 \hat{p} \cos\omega t, \tag{2.2}$$

characterizes the *(linear) polarization* of the electromagnetic wave. If $\theta = 0$ the light is polarized in the x direction, and if $\theta = \pi/2$ it is polarized in the y direction. Natural light is *unpolarized* because it is made up of an *incoherent* superposition (this important concept will be defined precisely in Chapter 4) of 50% light polarized along Ox and 50% light polarized along Oy.

We shall study polarization quantitatively using a *polarizer–analyzer ensemble*. We allow the light first to pass through a polarizer whose axis makes an angle θ with Ox, and then through a second polarizer, called an analyzer, whose axis makes an angle α with Ox (Fig. 2.1), and write

$$\hat{n} = (\cos\alpha, \sin\alpha). \tag{2.3}$$

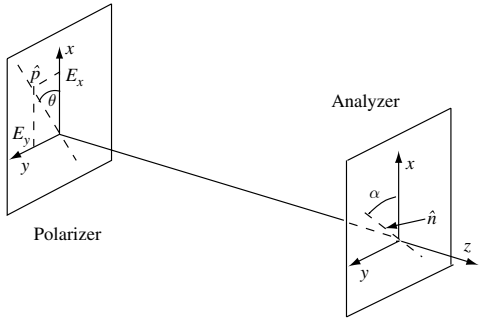

Figure 2.1 A polarizer–analyzer ensemble.

[1] Throughout this book, unit vectors of ordinary space $\mathbb{R}^3$ will be denoted by a hat: $\hat{p} = \vec{p}/p$, $\hat{n} = \vec{n}/n, \ldots$.

https://doi.org/10.1017/CBO9780511755361.004

At the exit from the analyzer the electric field $\vec{E}'$ is obtained by projecting the field (2.1) onto $\hat{n}$:

$$\begin{aligned}\vec{E}' &= (\vec{E}\cdot\hat{n})\hat{n} = E_0\cos\omega t(\hat{p}\cdot\hat{n})\hat{n}\\ &= E_0\cos\omega t(\cos\theta\cos\alpha+\sin\theta\sin\alpha)\hat{n}\\ &= E_0\cos\omega t\cos(\theta-\alpha)\hat{n}.\end{aligned} \tag{2.4}$$

From this we obtain the *Malus law* for the intensity at the exit from the analyzer:

$$I' = I\cos^2(\theta-\alpha). \tag{2.5}$$

Linear polarization is not the most general possible case. *Circular polarization* is obtained by choosing $\theta=\pi/4$ and shifting the phase of the y component by $\pm\pi/2$. For example, for right-handed circular polarization we have

$$\begin{aligned}E_x &= \frac{E_0}{\sqrt{2}}\cos\omega t,\\ E_y &= \frac{E_0}{\sqrt{2}}\cos\left(\omega t-\frac{\pi}{2}\right)=\frac{E_0}{\sqrt{2}}\sin\omega t.\end{aligned} \tag{2.6}$$

The electric field vector traces a circle of radius $|E_0|/\sqrt{2}$ in the xOy plane. The most general case is that of elliptical polarization, where the tip of the electric field vector traces an ellipse:

$$\begin{aligned}E_x &= E_0\cos\theta\cos(\omega t-\delta_x)=E_0\,\mathrm{Re}\left[\cos\theta\,\mathrm{e}^{-\mathrm{i}(\omega t-\delta_x)}\right]=E_0\,\mathrm{Re}\left(\lambda\,\mathrm{e}^{-\mathrm{i}\omega t}\right),\\ E_y &= E_0\sin\theta\cos(\omega t-\delta_y)=E_0\,\mathrm{Re}\left[\sin\theta\,\mathrm{e}^{-\mathrm{i}(\omega t-\delta_y)}\right]=E_0\,\mathrm{Re}\left(\mu\,\mathrm{e}^{-\mathrm{i}\omega t}\right).\end{aligned} \tag{2.7}$$

It will be important for what follows to note that *only the difference* $\delta=(\delta_y-\delta_x)$ *is physically relevant*. By a simple change of time origin we can, for example, choose $\delta_x=0$. To summarize, the most general polarization is described by a *complex* vector normalized to unity (or a *normalized vector*) in a two-dimensional space with components

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

and $|\lambda|^2+|\mu|^2=1$. Owing to the arbitrariness in the phase, a vector with components (λ',μ'),

$$\lambda'=\lambda\,\mathrm{e}^{\mathrm{i}\phi},\quad \mu'=\mu\,\mathrm{e}^{\mathrm{i}\phi},$$

https://doi.org/10.1017/CBO9780511755361.004

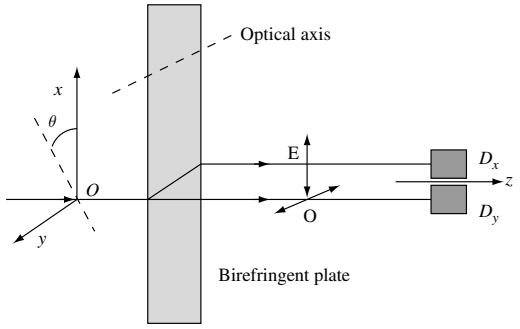

Figure 2.2 Decomposition of the polarization by a birefringent plate. The ordinary ray O is polarized horizontally and the extraordinary ray E is polarized vertically.

represents the same polarization as (λ, μ). It is more correct to say that the polarization is represented mathematically by a *ray*, that is, by a vector up to a phase.

Remarks

- A birefringent plate (Fig. 2.2) can be used to separate an incident beam into two orthogonal polarization states, and one can repeat the Malus experiment by checking that a suitably oriented polarizing filter absorbs one of the two polarizations while allowing the orthogonal one to pass through.
- Let us consider a crossed polarizer–analyzer ensemble, for example, with the polarizer aligned along Ox and the analyzer along Oy. No light is transmitted. However, if we introduce an intermediate polarizer whose axis makes an angle θ with Ox, part of the light is transmitted: the first projection gives a factor $\cos\theta$ and the second gives a factor $\sin\theta$, so that the intensity at the exit of the analyzer is

$$I' = I\cos^2\theta\sin^2\theta,$$

which vanishes only for $\theta = 0$ or $\theta = \pi/2$.

2.2 Photon polarization

Ever since the work of Einstein (1905), we have known that light is composed of photons or light particles. If the light intensity is reduced sufficiently, it should be possible to study the polarization of individual photons which can easily be detected using photodetectors, the modern version of which is the CCD (Charge Coupling Device) camera.[2] Let us suppose that $\mathcal{N}$ photons are detected in an

[2] A cell of the retina is sensitive to an isolated photon, but only a few percent of the photons entering the eye reach the retina.

experiment. When $\mathcal{N} \to \infty$ it should be possible to recover the results of wave optics which we have just stated above. For example, let us perform the following experiment (Fig. 2.2). A birefringent plate is used to separate a light beam whose polarization makes an angle θ with Ox into a beam polarized along Ox and a beam polarized along Oy, the intensities respectively being $I\cos^2\theta$ and $I\sin^2\theta$. We reduce the intensity such that the photons arrive one by one, and we place two photodetectors D_x and D_y behind the plate. Experiment shows that D_x and D_y are never triggered simultaneously,[3] i.e., an entire photon reaches *either* D_x *or* D_y: a photon is never split. On the other hand, experiment shows that the probability $\mathsf{p}_x(\mathsf{p}_y)$ that a photon is detected by $D_x(D_y)$ is $\cos^2\theta(\sin^2\theta)$. If $\mathcal{N}$ photons are detected in the experiment, we must have $\mathcal{N}_x(\mathcal{N}_y)$ photons detected by $D_x(D_y)$:

$$\mathcal{N}_x \simeq \mathcal{N}\cos^2\theta, \quad \mathcal{N}_y \simeq \mathcal{N}\sin^2\theta,$$

where $\simeq$ is used to indicate statistical fluctuations of order $\sqrt{\mathcal{N}}$. Since the light intensity is proportional to the number of photons, we recover the Malus law in the limit $\mathcal{N} \to \infty$. However, in spite of its simplicity this experiment raises two fundamental problems.

- **Problem 1** Is it possible to predict whether a given photon will trigger D_x or D_y? The response of quantum theory is NO, which profoundly shocked Einstein ("God does not play dice!"). Some physicists have tried to assume that quantum theory is incomplete, and that there are "hidden variables" whose knowledge would permit prediction of which detector a given photon reaches. If we make some very reasonable hypotheses to which we shall return in Chapter 4, we now know that such hidden variables are experimentally excluded. The probabilities of quantum theory are *intrinsic*; they are not related to imperfect knowledge of the physical situation, as is the case, for example, in the game of tossing a coin.
- **Problem 2** Let us recombine[4] the two beams from the first birefringent plate by using a second plate located symmetrically relative to the first (Fig. 2.3) and find the probability for a photon to cross the analyzer. A photon can choose path E with probability $\cos^2\theta$. Then it has probability $\cos^2\alpha$ of passing through the analyzer, or a total probability $\cos^2\theta\cos^2\alpha$. If path O is chosen, the probability of passing through the analyzer will be $\sin^2\theta\sin^2\alpha$. The total probability is obtained by adding the probabilities of the two possible paths:

$$\mathsf{p}'_{\text{tot}} = \cos^2\theta\cos^2\alpha + \sin^2\theta\sin^2\alpha. \tag{2.8}$$

[3] Except in the case of a "dark count," where a detector is triggered spontaneously.
[4] With some care, as the difference between the ordinary and extraordinary indices of refraction must be taken into account; *cf.* Le Bellac (2006), Exercise 3.1.

https://doi.org/10.1017/CBO9780511755361.004

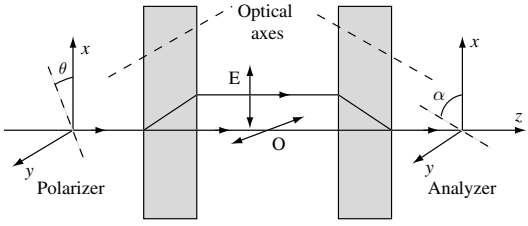

Figure 2.3 Decomposition and recombination of polarizations by means of birefringent plates. The photon can choose path E (extraordinary), where it is polarized along Ox, or path O (ordinary), where it is polarized along Oy.

This result is FALSE! In fact, we know from classical optics that the intensity is $I\cos^2(\theta-\alpha)$, and the correct result, confirmed by experiment, is

$$\mathsf{p}_{\text{tot}} = \cos^2(\theta-\alpha), \tag{2.9}$$

which is not at all the same thing!

In order to recover the results of wave optics it is necessary to introduce into quantum physics the fundamental notion of a *probability amplitude* $a(\alpha \to \beta)$. A probability amplitude is a complex number, the squared modulus of which gives the probability: $\mathsf{p}(\alpha \to \beta) = |a(\alpha \to \beta)|^2$. In the preceding example, the relevant probability amplitudes are

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

For example, $a(\theta \to x)$ is the probability amplitude that the photon polarized along the direction θ chooses the E path, where it is polarized along Ox. Then, a basic principle of quantum physics is that one must *add the amplitudes for indistinguishable paths*:

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

which allows us to recover (2.9):

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

The superposition of probability amplitudes in a_{tot} is the exact analog of the superposition of wave amplitudes: the laws for combining quantum amplitudes are exact copies of those of wave optics, and the results of the latter are recovered in the limit of a large number of photons. Let us suppose, however, that we have some way of knowing whether a photon has followed path E or path O (this is impossible in our case, but similar experiments to determine the path, termed

"which path experiments," have been performed using atoms). We can then divide the photons into two classes, those which have chosen path E and those which have chosen path O. For the former we could have blocked path O by a mask without changing anything, and the reverse for the latter photons. The result can obviously only be (2.8). If we manage to distinguish between the paths, the result will no longer be (2.9), because the paths are no longer indistinguishable!

Under experimental conditions where it is impossible in principle to distinguish between the paths, we can make one or the other statement:

- the photon is able to explore both paths at the same time, or
- (the author's preference) it makes no sense to ask the question "Which path?", because the experimental conditions do not permit it to be answered. We shall follow Asher Peres, who states "Unperformed experiments have no results!"

It should be noted that if the experiment allows us to distinguish between the two paths, the result is (2.8), even if we decide not to observe which path is followed. It is sufficient that the experimental conditions *in principle* allow the two paths to be distinguished, even when the current technology does not permit this to be done in practice.

We have examined a particular case of a quantum phenomenon, the photon polarization, but the results we have described have led us to the very core of quantum physics.

2.3 Mathematical formulation: the qubit

The photon polarization can be used to transmit information, for example, by an optical fiber. We can arbitrarily decide to associate the bit value 0 with a photon polarized along Ox and the bit value 1 with a photon polarized along Oy. In quantum information theory the people who exchange information are conventionally called Alice (A) and Bob (B). For example, Alice sends Bob a series of photons polarized as

$$yyxyxyyyx\cdots .$$

Bob analyzes the polarization of these photons using a birefringent plate as in Fig. 1.2 and deciphers the message sent by Alice:

$$110101110\cdots .$$

This is obviously not a very efficient way of exchanging messages. However, we shall see that this protocol forms the basis of quantum cryptography. An interesting question now is, what bit value can be associated with, for example, a photon polarized at 45°? According to the results of the preceding section, a

https://doi.org/10.1017/CBO9780511755361.004

photon polarized at 45° is a *linear superposition* of a photon polarized along Ox and a photon polarized along Oy. The photon polarization gives an example of a qubit, and a qubit is therefore a much richer object than an ordinary bit, which can take only the values 0 and 1. In a certain sense, a qubit can take all values intermediate between 0 and 1 and therefore contains an infinite amount of information! However, this optimistic statement is immediately deflated when we recall that measurement of a qubit can give only the result 0 or 1, no matter which basis is chosen: a photon either chooses the E path (value 0 of the bit) or the O path (value 1 of the bit) and this result holds whatever the orientation of the birefringent plate. Nevertheless, we can ask the question whether or not this "hidden information" contained in the linear superposition is valuable, and in Chapter 5 we shall see that under certain conditions this information can actually be exploited.

In order to take into account linear superpositions, it is natural to introduce a two-dimensional vector space $\mathcal{H}$ for the mathematical description of polarization. Any polarization state can be put into correspondence with a vector in this space. We can, for example, choose as orthogonal basis vectors of $\mathcal{H}$ the vectors $|x\rangle$ and $|y\rangle$ corresponding to linear polarizations along Ox and Oy. Any polarization state can be decomposed on this basis: [5]

$$|\Phi\rangle = \lambda|x\rangle + \mu|y\rangle. \tag{2.10}$$

We use the Dirac notation for the vectors of $\mathcal{H}$; see Box 2.1. There exists a very precise experimental procedure for constructing the state $|\Phi\rangle$; it is described in detail in Exercise 2.6.2. A linear polarization will be described using real coefficients λ and μ, but the description of a circular (2.6) or elliptical (2.7) polarization will require coefficients λ and μ which are complex. The space $\mathcal{H}$ is therefore a *complex vector space*, isomorphic to $\mathbb{C}^2$.

Probability amplitudes are associated with scalar products on this space. Let us take two vectors, $|\Phi\rangle$ given by (2.10) and $|\Psi\rangle$ given by

$$|\Psi\rangle = \nu|x\rangle + \sigma|y\rangle.$$

The *scalar product* of these vectors will be denoted $\langle\Psi|\Phi\rangle$, and by definition

$$\langle\Psi|\Phi\rangle = \nu^*\lambda + \sigma^*\mu = \langle\Phi|\Psi\rangle^*, \tag{2.11}$$

where c^* is the complex conjugate of c. This scalar product is therefore linear in $|\Phi\rangle$ and antilinear in $|\Psi\rangle$. It defines the *norm* $||\Phi||$ of the vector $|\Phi\rangle$:

$$||\Phi||^2 = \langle\Phi|\Phi\rangle = |\lambda|^2 + |\mu|^2. \tag{2.12}$$

[5] We use upper-case Greek letters for generic vectors of $\mathcal{H}$ in order to avoid confusion with the vectors representing linear polarizations such as $|\theta\rangle$, $|\alpha\rangle$, etc.

Box 2.1: Dirac notation

"Mathematicians tend to loathe the Dirac notation, because it prevents them from making distinctions they consider important. Physicists love the Dirac notation because they are always forgetting that such distinctions exist and the notation liberates them from having to remember" (Mermin (2003)). In our presentation here the Dirac notation reduces to a simple notational convention and avoids matters of principle.

Let $\mathcal{H}^{(N)}$ be a Hilbert space of finite dimension N on the complex numbers and let u, v, w be vectors of $\mathcal{H}^{(N)}$. The scalar product of two vectors v and w is denoted (v, w), following for the time being the mathematicians' notation. It satisfies [6]

$$(v, \lambda w + \mu w) = \lambda(v, w) + \mu(v, w'), \quad (v, w) = (w, v)^*.$$

Let $\{e_n\}$ be an orthonormal basis of $\mathcal{H}^{(N)}$, $n = 1, 2, \ldots, N$. In this basis the vectors (u, v, w) have the components

$$u_n = (e_n, u), \quad v_n = (e_n, v), \quad w_n = (e_n, w).$$

Let us consider a linear operator $A(v, w)$ defined by its matrix representation in the basis $\{e_n\}$:

$$A_{nm}(v, w) = v_n w_m^*.$$

The action of this operator on the vector u, $u \xrightarrow{A} u'$, is given in terms of components by

$$u'_n = \sum_m A_{nm}(v, w) u_m = \sum_m v_n w_m^* u_m = \sum_m v_n (w_m^* u_m) = v_n (w, u),$$

or in vector form

$$u' = A(v, w)u = v(w, u).$$

In Dirac notation, vectors are written as $|v\rangle$ and scalar products as $\langle w|v\rangle$:

$$v \to |v\rangle, \quad (w, v) \to \langle w|v\rangle.$$

With this notation the action of $A(v, w)$ is written as

$$\begin{aligned} |u'\rangle = |A(v, w)u\rangle &= |v\rangle\langle w|u\rangle \\ &= (|v\rangle\langle w|)|u\rangle, \end{aligned}$$

and the second line of this equation suggests the *notational convention*

$$A(v, w) = |v\rangle\langle w|.$$

[6] The convention of physicists differs from that of mathematicians in that for the latter the scalar product is antilinear in the second vector:

$$(v, \lambda w + \mu w') = \lambda^*(v, w) + \mu^*(v, w').$$

A case of particular importance is that where $v = w$ and v is a normalized vector. Then

$$A(v, v) = |v\rangle\langle v|, \quad |A(v, v)u\rangle = |v\rangle\langle v|u\rangle,$$

and $A(v, v)$ is the *projector* $\mathcal{P}_v$ onto the vector v, because $\langle v|u\rangle$ is the component of u along v. A familiar example is the projection in $\mathbb{R}^3$ of a vector $\vec{u}$ onto a unit vector $\hat{v}$:

$$\mathcal{P}_{\hat{v}}\vec{u} = \hat{v}(\vec{u} \cdot \hat{v}).$$

It is customary to use $|n\rangle$ to denote the vectors of an orthonormal basis: $e_n \to |n\rangle$, and the projector onto $|n\rangle$ then is

$$\mathcal{P}_n = |n\rangle\langle n|.$$

Let $\mathcal{H}^{(M)}$ be an M-dimensional ($M \leq N$) subspace of $\mathcal{H}^{(N)}$, and $|m\rangle$, $m = 1, 2, \ldots, M$ be an orthonormal basis in this subspace. The projector onto $\mathcal{H}^{(M)}$ then is

$$\mathcal{P}_{\mathcal{H}^{(M)}} = \sum_{m=1}^{M} |m\rangle\langle m|,$$

and if $M = N$ we obtain the decomposition of the identity, which physicists call the *completeness relation*:

$$\sum_{m=1}^{N} |m\rangle\langle m| = I,$$

where I is the identity operator. The matrix elements of a linear operator A are given by

$$A_{mn} = \langle m|An\rangle$$

and the completeness relation can be used, for example, to find immediately the matrix multiplication law:

$$(AB)_{mn} = \langle m|ABn\rangle = \langle m|AIBn\rangle = \sum_k \langle m|Ak\rangle\langle k|Bn\rangle = \sum_k A_{mk}B_{kn}.$$

The vectors $|x\rangle$ and $|y\rangle$ are orthogonal with respect to the scalar product (2.11) and they have unit norm:

$$\langle x|x\rangle = \langle y|y\rangle = 1, \qquad \langle x|y\rangle = 0.$$

The basis $\{|x\rangle, |y\rangle\}$ is therefore an *orthonormal basis* of $\mathcal{H}$. To the definition of a physical state we shall add the convenient, but not essential, normalization condition

$$||\Phi||^2 = |\lambda|^2 + |\mu|^2 = 1. \tag{2.13}$$

Polarization states will therefore be represented mathematically by normalized vectors (vectors of unit norm) in the space $\mathcal{H}$; they are called *state vectors* (of polarization). A vector space on which a positive-definite scalar product is defined is called a *Hilbert space*, and $\mathcal{H}$ is the *Hilbert space (of polarization states)*.

Now let us return to the probability amplitudes. A state linearly polarized along θ will be denoted $|\theta\rangle$ with

$$|\theta\rangle = \cos\theta\,|x\rangle + \sin\theta|y\rangle. \tag{2.14}$$

The vector $|\theta\rangle$ gives the mathematical description of the linear polarization state of a photon. The probability amplitude for a photon polarized along θ to pass through an analyzer oriented along α is, as we have seen above,

$$a(\theta \to \alpha) = \cos(\theta - \alpha) = \langle\alpha|\theta\rangle. \tag{2.15}$$

It is therefore given by the scalar product of the vectors $|\alpha\rangle$ and $|\theta\rangle$, and the probability of passing through the analyzer is given by the squared modulus of this amplitude (see (2.9)):

$$\mathsf{p}(\theta \to \alpha) = \cos^2(\theta - \alpha) = |\langle\alpha|\theta\rangle|^2. \tag{2.16}$$

A probability amplitude ("the amplitude of the probability for finding $|\Phi\rangle$ in $|\Psi\rangle$," where $|\Phi\rangle$ and $|\Psi\rangle$ represent general polarization states) will be defined in general as

$$a(\Phi \to \Psi) = \langle\Psi|\Phi\rangle, \tag{2.17}$$

and the corresponding probability will be given by

$$\mathsf{p}(\Phi \to \Psi) = |a(\Phi \to \Psi)|^2 = |\langle\Psi|\Phi\rangle|^2. \tag{2.18}$$

It is important to note that a state vector is actually defined only up to a multiplicative phase; for example, in (2.10) we can multiply λ and μ by the *same* phase factor

$$(\lambda, \mu) \equiv (\mathrm{e}^{\mathrm{i}\delta}\lambda, \mathrm{e}^{\mathrm{i}\delta}\mu),$$

because replacing $|\Phi\rangle$ by

$$|\Phi'\rangle = \mathrm{e}^{\mathrm{i}\delta}|\Phi\rangle$$

leaves the probabilities $|\langle\Psi|\Phi\rangle|^2$ unchanged whatever $|\Psi\rangle$, and these probabilities are the only quantities which can be measured. A multiplicative global phase is

https://doi.org/10.1017/CBO9780511755361.004

not physically relevant; the correspondence is therefore not between a physical state and a vector, but rather between a physical state and a *ray*, that is, a vector up to a phase.

Now we are ready to tackle the crucial question of *measurement* in quantum physics. Measurement is based on two notions, that of the preparation of a quantum state and that of a test. We again use the polarizer–analyzer ensemble and assume that the analyzer, which prepares the polarization state, is oriented along Ox. If the polarizer is also oriented along Ox, a photon leaving the polarizer passes through the analyzer with 100% probability, while if the polarizer is oriented along Oy, the probability is zero. The analyzer performs a *test* (of the polarization), and the result of the test is 1 or 0. The test then allows the polarization state of the photon to be determined. However, this is not the case in general. Let us assume that the polarizer is oriented in the direction θ or in the orthogonal direction $\theta_\perp$:

$$\begin{aligned} |\theta\rangle &= \cos\theta\,|x\rangle + \sin\theta\,|y\rangle, \\ |\theta_\perp\rangle &= -\sin\theta\,|x\rangle + \cos\theta\,|y\rangle. \end{aligned} \tag{2.19}$$

The states $|\theta\rangle$ and $|\theta_\perp\rangle$, like the states $|x\rangle$ and $|y\rangle$, form an orthonormal basis of $\mathcal{H}$. If, for example, the polarizer prepares the photon in the state $|\theta\rangle$ and the analyzer is oriented along Ox, then the probability of passing the test is $\cos^2\theta$. Two essential things should be noted:

- After the passage through the analyzer, the polarization state of the photon is no longer $|\theta\rangle$, but $|x\rangle$. It is often said that *the measurement perturbs the polarization state*. However, this statement is debatable: the measurement performed by the analyzer is a measurement of the physical property "polarization of the photon along Ox," but this polarization does not exist before the measurement because the photon is in the state $|\theta\rangle$, and that which does not exist cannot be perturbed! We shall illustrate this by another example at the end of this section.
- If the photon is elliptically, rather than linearly, polarized,

  $$\lambda = \cos\theta, \quad \mu = \sin\theta\,\mathrm{e}^{\mathrm{i}\delta}, \quad \delta \neq 0,$$

  the probability of passing the test is again $\cos^2\theta$: the test does not permit an unambiguous determination of the polarization. *Only if the probability of passing the test is 0 or 1 does the measurement permit the unambiguous determination of the initial polarization state. Therefore, unless one knows beforehand the basis in which it has been prepared, there is no test which permits the unambiguous determination of the polarization state of an isolated photon*. As explained in Exercise 2.6.1, determination of the polarization of a light wave, or of a large number of identically prepared photons, is possible provided one uses two different orientations of the analyzer.

There is thus a difference of principle between a measurement in classical physics and one in quantum physics. In classical physics *the physical quantity*

which is measured exists before the measurement: if radar is used to measure the speed of your car equal to 180 km/h on the highway, this speed existed before the police performed the measurement (thus giving them the right to issue a speeding ticket). On the contrary, in the measurement of the photon polarization $|\theta\rangle$ by an analyzer oriented along Ox, the fact that the test gives a polarization along Ox does not permit us to conclude that the tested photon actually had polarization along Ox before the measurement. Again taking the analogy to a car, we can imagine that as in (2.19) the car is in a linear superposition of two speed states, [7] for example,

$$|v\rangle = \sqrt{\frac{1}{3}}|120\,\text{km/h}\rangle + \sqrt{\frac{2}{3}}|180\,\text{km/h}\rangle.$$

The police will measure a speed of 120 km/h with probability 1/3 and a speed of 180 km/h with probability 2/3, but it would be incorrect to think that one of the two results existed before the measurement. Quantum logic is incompatible with classical logic!

2.4 Principles of quantum mechanics

The principles of quantum mechanics generalize the results we have obtained in the case of photon polarization.

- **Principle 1** The physical state of a quantum system is represented by a vector $|\Phi\rangle$ belonging to a Hilbert space $\mathcal{H}$ of, in general, infinite dimension. Fortunately, for the purposes of quantum information theory, we only need spaces of finite dimension. Unless explicitly stated otherwise, $|\Phi\rangle$ will be chosen to be a normalized vector: $||\Phi||^2 = 1$. $|\Phi\rangle$ is called the *state vector* of the quantum system.
- **Principle 2** If $|\Phi\rangle$ and $|\Psi\rangle$ represent two physical states, the probability amplitude $a(\Phi \to \Psi)$ of finding Φ in Ψ is given by the scalar product $\langle\Psi|\Phi\rangle$:

  $$a(\Phi \to \Psi) = \langle\Psi|\Phi\rangle,$$

  and the probability for Φ to pass the Ψ test is

  $$\mathsf{p}(\Phi \to \Psi) = |a(\Phi \to \Psi)|^2 = |\langle\Psi|\Phi\rangle|^2.$$

  We perform this test by first using a device to prepare the quantum system in the state $|\Phi\rangle$ (a polarizer), and then using as an analyzer a second device which would have prepared the system in the state $|\Psi\rangle$.

[7] Of course, no one knows how to realize such a superposition state for a car, but we do know very well how to construct a superposition of states with different speeds for an elementary particle or an atom.

https://doi.org/10.1017/CBO9780511755361.004

After the test the quantum system is in the state $|\Psi\rangle$, which from the mathematical point of view means that we have performed an orthogonal projection onto $|\Psi\rangle$. Let $\mathcal{P}_\Psi$ be the projector. Since [8]

$$|\mathcal{P}_\Psi\Phi\rangle \equiv \mathcal{P}_\Psi|\Phi\rangle = |\Psi\rangle\langle\Psi|\Phi\rangle = (|\Psi\rangle\langle\Psi|)|\Phi\rangle,$$

this projector can be written in the very convenient form (see Box 2.1)

$$\mathcal{P}_\Psi = |\Psi\rangle\langle\Psi|. \tag{2.20}$$

In summary, the mathematical operation corresponding to a measurement is a projection, and the corresponding measurement is called a *projective measurement*. However, the vector $\mathcal{P}_\Psi|\Phi\rangle$ is not in general normalized. We must then normalize it

$$\mathcal{P}_\Psi|\Phi\rangle \to |\Phi'\rangle = \frac{\mathcal{P}_\Psi|\Phi\rangle}{\langle\Phi|\mathcal{P}_\Psi\Phi\rangle}.$$

In the orthodox interpretation of quantum mechanics, the projection of a state vector followed by its normalization is called "state-vector collapse" or, for historical reasons, "wave-packet collapse." The idea of state-vector collapse is a convenient fiction of the orthodox interpretation which avoids having to ask questions about the measurement process, and it is often treated as a supplementary basic principle of quantum mechanics. However, we can perfectly well bypass this principle if we take into account the full complexity of the measurement process. An example will be given in Chapter 5, Box 5.2.

Let us now turn to the mathematical description of the physical properties of a quantum system, first by returning to polarization. In the basis $\{|x\rangle, |y\rangle\}$ the projectors $\mathcal{P}_x$ and $\mathcal{P}_y$ onto these basis states are

$$\mathcal{P}_x = |x\rangle\langle x| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathcal{P}_y = |y\rangle\langle y| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

We note that the identity operator I can be written as the sum of the two projectors $\mathcal{P}_x$ and $\mathcal{P}_y$:

$$\mathcal{P}_x + \mathcal{P}_y = |x\rangle\langle x| + |y\rangle\langle y| = I.$$

This is a special case of the *completeness relation* (Box 2.1), which can be generalized to an orthonormal basis of a Hilbert space $\mathcal{H}$ of dimension N:

$$\sum_{i=1}^{N} |i\rangle\langle i| = I, \quad \langle i|j\rangle = \delta_{ij}.$$

[8] The action of an operator M on a vector $|\Phi\rangle$ will be written either as $M|\Phi\rangle$ or as $|M\Phi\rangle$.

The projectors $\mathcal{P}_x$ and $\mathcal{P}_y$ commute:

$$[\mathcal{P}_x, \mathcal{P}_y] \equiv \mathcal{P}_x\mathcal{P}_y - \mathcal{P}_y\mathcal{P}_x = 0,$$

where we denote $[A, B] := AB - BA$ the *commutator* of two operators A and B. The *tests* $|x\rangle$ and $|y\rangle$ are termed *compatible*. On the contrary, the projectors onto the states $|\theta\rangle$ and $|\theta_\perp\rangle$ (2.19),

$$\mathcal{P}_\theta = |\theta\rangle\langle\theta| = \begin{pmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{pmatrix},$$

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

do not commute with $\mathcal{P}_x$ and $\mathcal{P}_y$, as can be verified immediately by explicit calculation:

$$[\mathcal{P}_x, \mathcal{P}_\theta] = \begin{pmatrix} 0 & \sin\theta\cos\theta \\ -\sin\theta\cos\theta & 0 \end{pmatrix}.$$

The *tests* $|x\rangle$ and $|\theta\rangle$ are termed *incompatible*. The projectors $\mathcal{P}_x, \ldots, \mathcal{P}_{\theta_\perp}$ represent mathematically the physical properties of the quantum system when the photon is polarized along the $x, \ldots, \theta_\perp$ axes. *It is not possible to measure incompatible properties of a quantum system simultaneously*.

In the general case of a Hilbert space of states $\mathcal{H}^{(N)}$ with dimension N, to an orthonormal basis $|n\rangle$, $n = 1, \ldots, N$, of this space will be associated a set of N compatible tests $|n\rangle$. If the quantum system is in a state $|\Phi\rangle$, the probability that it passes the test is $\mathsf{p}_n = \langle n|\Phi\rangle|^2$ (**Principle 2**) and $\sum_n \mathsf{p}_n = 1$. The tests $|n\rangle$ form a *maximal test*. To a different orthonormal basis of $\mathcal{H}^{(N)}$ will correspond another maximal test incompatible with the preceding one. Now, one may ask the following question: can one define in a Hilbert space $\mathcal{H}^{(N)}$ bases $\{|n\rangle\}$ and $\{|\alpha\rangle\}$, where $n, \alpha = 1, 2, \ldots, N$, which are maximally incompatible? The answer to this question is positive. Two bases are maximally incompatible if they are *complementary*, which means by definition that $|\langle\alpha|n\rangle|^2$ is independent of n and α

$$|\langle\alpha|n\rangle|^2 = \frac{1}{N}. \tag{2.21}$$

For example, the bases $\{|x\rangle, |y\rangle\}$ and $\{|\theta = \pi/4\rangle, |\theta = -\pi/4\rangle\}$ are complementary. Any linear polarization basis is complementary to a circular polarization basis $\{|R\rangle, |L\rangle\}$ defined in Exercise 2.6.3. One way to obtain complementary bases in $\mathcal{H}^{(N)}$ is to use discrete Fourier transforms (see Section 5.7)

$$|\alpha\rangle = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \mathrm{e}^{2\mathrm{i}\pi\alpha n}|n\rangle.$$

Let us explain the physical meaning of complementary bases with the following example: suppose you want to test a large number of quantum systems all prepared in a state of the basis $\{|n\rangle\}$, but you do not know which one. If you test the system using the basis $\{|n\rangle\}$, one of the results, say m, will come out with 100% probability, so that your measurement gives you maximum knowledge of the state. If, on the contrary, you test the preparation using the basis $\{|\alpha\rangle\}$, then you will get all the possible outcomes with probability $1/N$, and you will get minimum knowledge of the preparation. The concept of complementary bases will be very useful for understanding the principles of quantum cryptography.

For later developments it will be useful to note that knowledge of the probabilities of passing a test $\mathcal{T}$ permits definition of the *expectation value* $\langle\mathcal{T}\rangle$:

$$\langle\mathcal{T}\rangle = 1\times \mathsf{p}(\mathcal{T}=1)+0\times \mathsf{p}(\mathcal{T}=0) \quad [=\mathsf{p}(\mathcal{T}=1)].$$

For example, if the test $\mathcal{T}$ is represented by the procedure $|\Psi\rangle$ and it is applied to a state $|\Phi\rangle$, then

$$\mathsf{p}(\Psi) = |\langle\Psi|\Phi\rangle|^2 = \langle\Phi|\Psi\rangle\langle\Psi|\Phi\rangle = \langle\Phi(|\Psi\rangle\langle\Psi)|\Phi\rangle = \langle\Phi|\mathcal{P}_\Psi\Phi\rangle. \qquad (2.22)$$

In quantum physics it is standard to refer to the quantity

$$\boxed{\langle\Phi|M\Phi\rangle \equiv \langle M\rangle_\Phi} \qquad (2.23)$$

as the *expectation value of the operator M in the state* $|\Phi\rangle$. The test $\mathcal{T}=|\Psi\rangle$ can therefore be associated with a projector $\mathcal{P}_\Psi$ whose expectation value in the state $|\Phi\rangle$ gives, according to (2.22), the probability of passing the test.

The generalization of this observation permits us to construct the physical properties of a quantum system using projectors. Let us give an example, again from the case of polarization. We assume that we have constructed (in a completely arbitrary way) a physical property $\mathcal{M}$ of a photon as follows: $\mathcal{M}$ is $+1$ if the photon is polarized along Ox and $\mathcal{M}$ is -1 if the photon is polarized along Oy. With the physical property $\mathcal{M}$ we can associate a Hermitian operator M,

$$M = \mathcal{P}_x - \mathcal{P}_y,$$

satisfying

$$M|x\rangle = +|x\rangle, \quad M|y\rangle = -|y\rangle.$$

The expectation value of M is, by definition,

$$\langle M\rangle = 1\times \mathsf{p}(M=1)+(-1)\times \mathsf{p}(M=-1).$$

Let us assume that the photon is in the state $|\theta\rangle$; then the expectation value $\langle M\rangle_\theta$ in the state $|\theta\rangle$ is

$$\langle M\rangle_\theta = \langle\theta|\mathcal{P}_x\theta\rangle - \langle\theta|\mathcal{P}_y\theta\rangle = \cos^2\theta - \sin^2\theta = \cos 2\theta.$$

The operator M thus constructed is a Hermitian operator ($M = M^\dagger$ or $M_{ij} = M^*_{ji}$), and in general *physical properties* in quantum mechanics are represented mathematically by Hermitian operators, often called *observables*. We have constructed M starting from projectors, but reciprocally we can construct projectors starting from a Hermitian operator M owing to the *spectral decomposition theorem*, which we state without proof.

Theorem Let M be a Hermitian operator. Then M can be written as a function of a set of projectors $\mathcal{P}_n$ satisfying

$$M = \sum_n a_n \mathcal{P}_n, \tag{2.24}$$

$$\mathcal{P}_n\mathcal{P}_m = \mathcal{P}_n\delta_{mn}, \qquad \sum_n \mathcal{P}_n = I, \tag{2.25}$$

where the real coefficients a_n are the eigenvalues of M. The projectors $\mathcal{P}_n$ are orthogonal to each other (but in general they project onto a subspace of $\mathcal{H}$ and not onto a single vector of $\mathcal{H}$), and their sum is the identity operator.

Let us summarize the results on the physical properties. The physical properties of a quantum system are represented mathematically by Hermitian operators and the measurement of a physical property $\mathcal{M}$ has as its result one of the eigenvalues a_n of the operator M

$$M|n\rangle = a_n|n\rangle.$$

In order to simplify the discussion, we assume that the eigenvalues of M are nondegenerate, so that the spectral decomposition (2.24) and (2.25) becomes

$$M = \sum_{n=1}^{N} a_n|n\rangle\langle n| \quad I = \sum_{n=1}^{N} |n\rangle\langle n|, \tag{2.26}$$

where N is the dimension of the Hilbert space of states. If the quantum system is in the eigenstate $|n\rangle$, the value taken by $\mathcal{M}$ is *exactly* a_n. If the quantum system is in the state $|\Phi\rangle$, then the probability of finding it in $|n\rangle$ is, from (2.18)

$$\mathsf{p}_n = |\langle n|\Phi\rangle|^2 = \langle\Phi|n\rangle\langle n|\Phi\rangle.$$

https://doi.org/10.1017/CBO9780511755361.004

If one measures the value a_n of M, then the state vector after measurement is $|n\rangle$: this is the state vector collapse. The expectation value of $\mathcal{M}$ is by definition

$$\sum_{n=1}^{N} \mathsf{p}_n a_n = \sum_{n=1}^{N} \langle\Phi|n\rangle a_n \langle n|\Phi\rangle = \langle\Phi|M\Phi\rangle = \langle M\rangle_\Phi, \tag{2.27}$$

which justifies the definition (2.23). This expectation value has the following physical interpretation: in an experiment conducted with a large number $\mathcal{N}$ of quantum systems all prepared in the same state $|\Phi\rangle$, the average value of the measurements of M is $\langle M\rangle_\Phi$

$$\langle M\rangle_\Phi = \lim_{\mathcal{N}\to\infty} \frac{1}{\mathcal{N}} (M_1 + \cdots + M_\mathcal{N}) \tag{2.28}$$

where M_i is the result of the measurement number i, which is necessarily one of the eigenvalues a_n of M. We leave to the reader the generalization of the preceding results to the case where M has degenerate eigenvalues.

Box 2.2: A quantum random-number generator

It is often necessary to generate random numbers, for example, for use in Monte Carlo simulations. All computers contain a program for random number generation. However, these numbers are generated by an algorithm, and they are not actually random, but only *pseudo-random*. A simple algorithm (too simple to be reliable!) consists of, for example, calculating

$$I_{n+1} \equiv aI_n + b \bmod \mathrm{M}, \quad 0 \le I_n \le M-1,$$

where a and b are given integers and M is an integer, $M \gg 1$. The series $I'_n = I_n/M$ is a series of pseudo-random numbers in the interval $[0, 1]$. In some cases the inevitable regularities in a series of pseudo-random numbers can lead to errors in numerical simulations. Quantum properties can be used experimentally to realize generators of numbers which are truly random and not pseudo-random; as we shall see in the following section, truly random numbers are essential for quantum cryptography. One of the simplest devices uses a semi-transparent plate or beam-splitter. If a light ray falls on a beam-splitter, part of the light is transmitted and part is reflected. This can be arranged such that the proportions are 50%/50%. If the intensity is then decreased such that the photons arrive one by one at the plate, these photons can be either reflected and detected by D_1, or transmitted and detected by D_2 (Fig. 2.4). There is no correlation between the detections, and so this amounts to a true, unbiased coin toss. A prototype based on this principle has been realized by the quantum optics group in Geneva. It generates random numbers at a rate of 10^5 numbers per second, and the absence of any bias (equivalently, correlations between numbers supposed to be random) has been verified using standard programs.

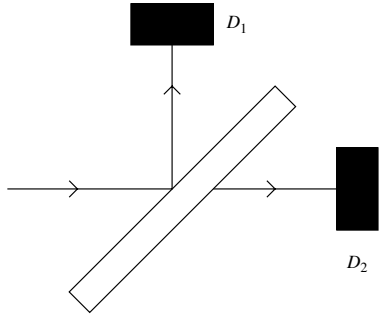

Figure 2.4 A semi-transparent plate and photon detection.

2.5 Quantum cryptography

Quantum cryptography is a recent invention based on the incompatibility of two different bases of linear polarization states. Ordinary cryptography is based on an encryption key known only to the sender and receiver and is called *secret-key cryptography*. It is in principle very secure,[9] but the sender and receiver must have a way of exchanging the key without it being intercepted by a spy. The key must be changed frequently, because a set of messages encoded with the same key can reveal regularities which permit decipherment by a third person. The transmission of a secret key is a risky process, and for this reason it is now preferred to use systems based on a different principle, the so-called *public-key* systems. In these the key is announced publicly, for example, via the Internet. A public-key system currently in use[10] is based on the difficulty of factoring a very large number N into primes, whereas the reverse operation can be done immediately: even without the help of a pocket calculator one can find $137 \times 53 = 7261$ in a few seconds, but given 7261 it would take a some time to factor it into primes. Using the best current algorithms, the time needed for a computer to factor a number N into primes grows with N as $\simeq \exp[1.9(\ln N)^{1/3}(\ln \ln N)^{2/3}]$. The current record is 176 digits, and it takes several months for a PC cluster to factorize such a number. In public-key encryption the receiver, conventionally named Bob, publicly announces to the sender, conventionally named Alice, a very large number $N = pq$ which is the product of two prime numbers p and q, along with another number c (see Box 2.3). These two numbers N and c are sufficient for Alice to encode the message, but the numbers p and q are needed to decipher it. Of course, a spy

[9] 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!
[10] Called RSA, as it was invented by Rivest, Shamir, and Adleman in 1977.

Box 2.3: RSA encryption (see also Box 5.3)

Bob chooses two primes p and q, $N = pq$, and a number c having no common divisor with the product $(p-1)(q-1)$. He calculates d, the inverse of c for mod $(p-1)(q-1)$ multiplication:

$$cd \equiv 1 \bmod (p-1)(q-1).$$

By a non-secure path he sends Alice the numbers N and c (but not p and q separately!). Alice wants to send Bob an encoded message, which must be represented by a number $a < N$ (if the message is too long, Alice can split it into several sub-messages). She then calculates (Fig. 2.5)

$$b \equiv a^c \bmod N$$

and sends b to Bob, always by a non-secure path, because a spy who knows only N, c, and b cannot deduce the original message a. When Bob receives the message he calculates

$$b^d \bmod N = a.$$

The fact that the result is precisely a, that is, the original message of Alice, is a result from number theory (see Box 5.3 for a proof of this result). To summarize, the numbers N, c, and b are sent publicly, by a non-secure path.

Example

$$p = 3, \quad q = 7, \quad N = 21, \quad (p-1)(q-1) = 12.$$

The number $c = 5$ has no common factor with 12, and its inverse with respect to mod 12 multiplication is $d = 5$ because $5 \times 5 = 24 + 1$. Alice chooses $a = 4$ for her message. She calculates

$$4^5 = 1024 = 21 \times 48 + 16, \quad 4^5 \equiv 16 \bmod 21.$$

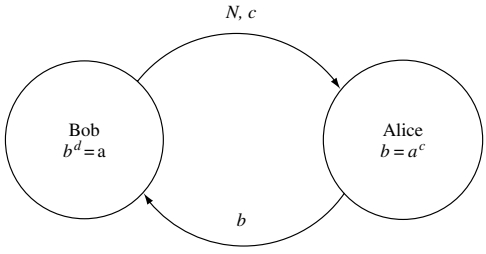

Figure 2.5 RSA encryption scheme. Bob chooses $N = pq$ and c. Alice encrypts her message a using $b = a^c$ and Bob decrypts it using $b^d = a$.

Alice then sends Bob the message 16. Bob calculates

$$b^5 = 16^5 = 49.932 \times 21 + 4, \quad 16^5 \equiv 4 \bmod 21,$$

thus recovering the original message $a = 4$. The above calculation of $16^5 \bmod 21$, for example, has not been done very cleverly. Instead, we can calculate $16^2 \bmod 21 = 4$, and then $16^3 \bmod 21$ as $4 \times 16 \bmod 21 = 1$, from which without further calculation we find $16^5 \bmod 21 = 4$. This method can be used to manipulate only numbers which are not very large compared to N.

(conventionally named Eve) possessing a sufficiently powerful computer and enough time will eventually crack the code, but one can in general count on the message being kept secret for a limited period of time. However, it is not impossible that one day we will possess very powerful algorithms for decomposing a number into primes, and moreover, if quantum computers ever see the light of day such factorization will become quite simple, at least in principle. Happily, thanks to quantum mechanics we are nearly at the point of being able to thwart the efforts of spies!

"Quantum cryptography" is a catchy phrase, but it is somewhat inaccurate. A better terminology is *quantum key distribution* (QKD). In fact, there is no encryption of a message using quantum physics; the latter is used only to ensure that the transmission of a key is not intercepted by a spy. As we have already explained, 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 choose to associate the value 0 with the polarization $|x\rangle$ and the value 1 with the polarization $|y\rangle$, so that each photon will carry a bit of information. Any message, encrypted or not, can be written in binary language as a series of 0s and 1s. The message 0110001 will be encoded by Alice by the photon sequence $xyyxxxy$, which she will send to Bob via, for example, an optical fiber. Bob will use a birefringent plate to separate the photons of vertical and horizontal polarization as in Fig. 2.2, and two detectors placed behind the plate will tell him whether the photon was polarized horizontally or vertically, so that he can reconstruct the message. If the message were just an ordinary one, there would certainly be much easier and more efficient ways of sending it! Let us simply note that if Eve taps into the fiber, detects the photons, and then resends to Bob photons of polarization identical to the ones sent by Alice, then Bob has no way of knowing that the transmission has been intercepted. The same would be true for any apparatus functioning in a classical manner (that is, not using the superposition principle): if the spy takes sufficient precautions, the spying is undetectable.

This is where quantum mechanics and the superposition principle come to the aid of Alice and Bob, by allowing them to be sure that their message has not

been intercepted. The message need not be long (the transmission scheme based on polarization is not very efficient). In general, one wishes to transmit a key which permits the encryption of a later message, a key which can be replaced whenever desired. Alice sends Bob photons of four types, polarized along $Ox(\updownarrow)$ and $Oy(\leftrightarrow)$, as before, and polarized along axes rotated by $\pm 45°$: $Ox'(\searrow)$ and $Oy'(\nearrow)$, respectively corresponding to bit values 0 and 1 (Fig. 2.6). Note that the two bases are complementary, or maximally incompatible. Similarly, Bob analyzes the photons sent by Alice using analyzers which can be oriented in four directions, vertical/horizontal and $\pm 45°$. One possibility is to use a birefringent crystal randomly oriented either vertically or at 45° with respect to the vertical and to detect the photons leaving the crystal as in Fig. 2.3. However, instead of rotating the crystal+detectors ensemble, it is easier to use a Pockels cell, which allows a given polarization to be transformed into an arbitrary polarization while maintaining the crystal+detectors ensemble fixed. An example is given in Fig. 2.7. Bob records a 0 if the photon has polarization $\updownarrow$ or $\nearrow$ and 1 if it has polarization $\leftrightarrow$ or $\nwarrow$. After recording a sufficient number of photons, Bob publicly announces the sequence of analyzers he has used, but not his results. Alice compares her sequence of polarizers to Bob's analyzers and publicly gives him the list of polarizers compatible with his analyzers. The bits corresponding to incompatible analyzers and polarizers are rejected $(-)$, and then Alice and Bob are certain that the values of the other bits are the same. These are the bits which will be used 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! The protocol we have described is called BB84, from the names of its inventors Bennett and Brassard.

We still need to be sure that the message has not been intercepted and that the key it contains can be used without risk. Alice and Bob choose at random a subset of their key and compare their choices publicly. The consequence of interception of the photons by Eve will be a reduction of the correlation between the values of their bits. Let us suppose, for example, that Alice sends a photon

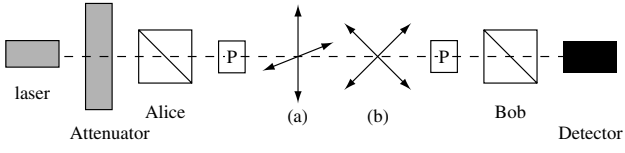

Figure 2.6 Schematical depiction of the BB84 protocol. A laser beam is attenuated such that it sends individual photons. A birefringent plate selects the polarization, which can be rotated by means of Pockels cells P. The photons are either vertically/horizontally polarized (a) or polarized at $\pm 45°$ (b).

Alice's polarizers	↕	↔	↗	↕	↗	↗	↘	↕	↘
Sequence of bits	0	1	1	0	1	1	0	0	0
Bob's polarizers	✛	✕	✛	✛	✕	✕	✛	✛	✕
Bob's measurements	0	0	1	0	1	1	0	0	0
Retained bits	0	–	–	0	1	1	–	0	0

Figure 2.7 Quantum cryptography: transmission of polarized photons between Bob and Alice.

polarized along Ox. If Eve intercepts it using a polarizer oriented along Ox', and if the photon passes through her analyzer, she does not know that this photon was initially polarized along Ox. She then resends to Bob a photon polarized in the direction Ox', and in 50% of cases Bob will not obtain the correct result. Since Eve has one chance in two of orienting her analyzer in the right direction, Alice and Bob will record a difference in 25% of the cases and will conclude that the message has been intercepted. It is the use of two complementary bases which allows the maximal security (Exercise 2.6.4). To summarize: the security of the protocol depends on the fact that Eve cannot find out the polarization state of a photon unless she knows beforehand the basis in which it was prepared.

Of course, this discussion has been simplified considerably. It does not take into account the possibility of errors which must be corrected thanks to a classical error-correcting code, while another classical process, called privacy amplification, ensures the secrecy of the key. Moreover, the scheme should be realized using single photons and not packets of coherent states like those produced by the attenuated laser of Fig. 2.6, which are less secure, but are used for practical reasons. [11] The *quantum bit error rate* (QBER) is simply the probability that Bob measures the wrong value of the polarization when he knows Alice's basis, for example the probability that he measures a y polarization while a photon polarized along Ox was sent by Alice. It can be shown that the QBER must be less than 11% if Alice and Bob want to obtain a reliable key. The QEBR q must obey

$$q \leq q_0 \quad q_0 \log q_0 + (1 - q_0) \log(1 - q_0) = \frac{1}{2},$$

[11] For example, an attenuated laser pulse contains typically 0.1 photon on the average. It can then be shown that a nonempty pulse has a 5% probability of containing two photons, a fact which can be exploited by Eve. In the case of transmission of isolated photons, the quantum no-cloning theorem (Section 4.3) guarantees that it is impossible for Eve to trick Bob, even if the error rate can be decreased to less than 25% by using a more sophisticated interception technique.

where log is a base 2 logarithm. If Eve is limited on attacking individual qubits, [12] the QBER must be less than 15%:

$$q \leq q_0' = \frac{1 - 1/\sqrt{2}}{2}.$$

A prototype has recently been realized for transmissions of several kilometers through air. When an optical fiber is used it is difficult to control the direction of the polarization over large distances, and so in that case a different physical support is needed to implement the BB84 protocol. Transmission over about a hundred kilometers has been achieved using optical fibers, and two versions of the device are available on the market.

2.6 Exercises

2.6.1 Determination of the polarization of a light wave

1. The polarization of a light wave is described by two complex parameters

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

satisfying $|\lambda|^2 + |\mu|^2 = 1$. More explicitly, the electric field is

$$E_x(t) = E_0 \cos\theta \cos(\omega t - \delta_x) = E_0\, \mathrm{Re}\left(\cos\theta\, e^{i\delta_x} e^{-i\omega t}\right),$$

$$E_y(t) = E_0 \sin\theta \cos(\omega t - \delta_y) = E_0\, \mathrm{Re}\left(\sin\theta\, e^{i\delta_y} e^{-i\omega t}\right).$$

Determine the axes of the ellipse traced by the tip of the electric field vector and the direction in which it is traced.

2. This light wave is made to pass through a polarizing filter whose axis is parallel to Ox. Show that measurement of the intensity at the exit of the filter allows θ to be determined.

3. Now the filter is oriented such that its axis makes an angle of $\pi/4$ with Ox. What is the reduction of the intensity at the exit from the filter? Show that this second measurement permits determination of the phase difference $\delta = \delta_y - \delta_x$.

2.6.2 The (λ, μ) polarizer

1. In (2.7) we use complex notation:

$$E_x(t) = E_{0x} \cos(\omega t - \delta_x) = \mathrm{Re}\left(E_{0x}\, e^{i\delta_x} e^{-i\omega t}\right) = \mathrm{Re}\left(\mathcal{E}_x\, e^{-i\omega t}\right),$$

$$E_y(t) = E_{0y} \cos(\omega t - \delta_y) = \mathrm{Re}\left(E_{0y}\, e^{i\delta_y} e^{-i\omega t}\right) = \mathrm{Re}\left(\mathcal{E}_y\, e^{-i\omega t}\right).$$

[12] That is, she is not allowed to store many qubits, which would permit coherent attacks on many qubits.

https://doi.org/10.1017/CBO9780511755361.004

Let the two numbers λ real and μ complex be parametrized as

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

A (λ, μ) polarizer is constructed of three elements.

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

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

- A linear polarizer which projects onto the unit vector $\hat{n}_\theta = (\cos\theta, \sin\theta)$:

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

- A second birefringent plate which leaves $\mathcal{E}_x^{(2)}$ unchanged and changes 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 all three operations is represented as $\vec{\mathcal{E}} \to \vec{\mathcal{E}}'$. Calculate the components $\mathcal{E}'_x$ and $\mathcal{E}'_y$ as functions of $\mathcal{E}_x$ and $\mathcal{E}_y$.

2. Vectors of $\mathcal{H}$ which are not normalized, $|\mathcal{E}\rangle$ and $|\mathcal{E}'\rangle$, are defined as

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

Show that the operation $|\mathcal{E}\rangle \to |\mathcal{E}'\rangle$ is a projection:

$$|\mathcal{E}'\rangle = \mathcal{P}_\Phi\,|\mathcal{E}\rangle,$$

where $\mathcal{P}_\Phi$ is the projector onto the vector

$$|\Phi\rangle = \lambda|x\rangle + \mu|y\rangle.$$

3. Show that a photon with state vector $|\Phi\rangle$ is transmitted by the (λ, μ) polarizer with unit probability, and that a photon of state vector

$$|\Phi_\perp\rangle = -\mu^*|x\rangle + \lambda^*|y\rangle$$

is stopped by this polarizer.

2.6.3 Circular polarization and the rotation operator

1. Justify the following expressions for the states $|R\rangle$ and $|L\rangle$ respectively representing right- and left-handed polarized photons:

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

where $|x\rangle$ and $|y\rangle$ are the state vectors of photons linearly polarized along Ox and Oy. Hint: what is the electric field of a circularly polarized light wave? Write down the matrix form of the projectors $\mathcal{P}_R$ and $\mathcal{P}_L$ onto the states $|R\rangle$ and $|L\rangle$ in the basis $\{|x\rangle, |y\rangle\}$.

2. We define the states $|\theta\rangle$ and $|\theta_\perp\rangle$ (2.19) representing photons linearly polarized along directions making an angle θ with Ox and Oy, respectively, and also the states

$$|R'\rangle = \frac{1}{\sqrt{2}}(|\theta\rangle + \mathrm{i}|\theta_\perp\rangle), \quad |L'\rangle = \frac{1}{\sqrt{2}}(|\theta\rangle - \mathrm{i}|\theta_\perp\rangle).$$

How are $|R'\rangle$ and $L'\rangle$ related to $|R\rangle$ and $|L\rangle$? Do these state vectors represent physical states different from $|R\rangle$ and $|L\rangle$? If not, then why not?

3. We construct the Hermitian operator

$$\Sigma = \mathcal{P}_R - \mathcal{P}_L.$$

What is the action of Σ on the vectors $|R\rangle$ and $|L\rangle$? Determine the action of $\exp(-\mathrm{i}\theta\Sigma)$ on these vectors.

4. Write the matrix representing Σ in the basis $\{|x\rangle, |y\rangle\}$. Show that $\Sigma^2 = I$ and recover $\exp(-\mathrm{i}\theta\Sigma)$. By comparing with question **2**, give the physical interpretation of the operator $\exp(-\mathrm{i}\theta\Sigma)$.

2.6.4 An optimal strategy for Eve?

1. Let us suppose that Eve analyzes the polarization of a photon sent by Alice using an analyzer oriented as ↕. If Alice orients her polarizer as ↕, the probability that Eve measures the value of the qubit as $+1$ is 100% when Alice sends a qubit $+1$, but only 50% when Alice uses a ↘ polarizer. The probability that Eve measures $+1$ when Alice sends $+1$ then is

$$\mathsf{p} = \frac{1}{2} + \frac{1}{2}\left(\frac{1}{2}\right) = \frac{3}{4}.$$

Let us suppose that Eve orients her analyzer in a direction making an angle ϕ with Ox. Show that the probability $\mathsf{p}(\phi)$ for Eve to measure $+1$ when Alice sends $+1$ is now

$$\mathsf{p}(\phi) = \frac{1}{4}\left(2 + \cos 2\phi + \sin 2\phi\right).$$

Show that for the optimal choice $\phi = \phi_0 = \pi/8$

$$\mathsf{p}(\phi_0) \simeq 0.854,$$

a larger value than before. Would it have been possible to predict without calculation that the optimal value must be $\phi = \phi_0 = \pi/8$? However, as explained in Section 7.2, the information gain of Eve is less than with the naive strategy.

https://doi.org/10.1017/CBO9780511755361.004

2. Suppose that instead of using a basis $|\pm\pi/4\rangle$, Alice uses a $\{|\theta\rangle, |\theta_\perp\rangle\}$ basis. Show that the probability that Eve makes a wrong guess is now

$$\mathsf{p} = \frac{1}{4}\sin^2(2\theta).$$

Thus, the use of complementary bases maximizes Eve's error rate.

2.6.5 Heisenberg inequalities

1. Let us take two Hermitian operators A and B. Show that their *commutator* $[A, B]$ is anti-Hermitian,

$$[A, B] := AB - BA = \mathrm{i}C,$$

where C is Hermitian: $C = C^\dagger$.

2. The expectation values of A and B are defined as

$$\langle A\rangle_\varphi = \langle\varphi|A\varphi\rangle, \quad \langle B\rangle_\varphi = \langle\varphi|B\varphi\rangle,$$

and the *dispersions* $\Delta_\varphi A$ and $\Delta_\varphi B$ *in the state* $|\varphi\rangle$ as

$$(\Delta_\varphi A)^2 = \langle A^2\rangle_\varphi - (\langle A\rangle_\varphi)^2 = \langle (A - \langle A\rangle_\varphi I)^2\rangle_\varphi,$$
$$(\Delta_\varphi B)^2 = \langle B^2\rangle_\varphi - (\langle B\rangle_\varphi)^2 = \langle (B - \langle B\rangle_\varphi I)^2\rangle_\varphi.$$

Finally, we define Hermitian operators of zero expectation value (which are *a priori specific to the state* $|\varphi\rangle$) as

$$A_0 = A - \langle A\rangle_\varphi I, \quad B_0 = B - \langle B\rangle_\varphi I.$$

What is their commutator? The norm of the vector

$$(A_0 + \mathrm{i}\lambda B_0)|\varphi\rangle,$$

where λ is chosen to be real, must be positive:

$$||(A_0 + \mathrm{i}\lambda B_0)|\varphi\rangle|| \geq 0.$$

Derive the Heisenberg inequality

$$(\Delta_\varphi A)\,(\Delta_\varphi B) \geq \frac{1}{2}\left|\langle C\rangle_\varphi\right|.$$

Care must be taken in interpreting this inequality. It implies that if a large number of quantum systems are prepared in the state $|\varphi\rangle$, and if their expectation values and dispersions $\{\langle A\rangle_\varphi, \Delta_\varphi A\}$, $\{\langle B\rangle_\varphi, \Delta_\varphi B\}$, and $\langle C\rangle_\varphi$ are measured in *independent* experiments, then these expectation values will obey the Heisenberg inequality. In contrast to what is sometimes found in the literature, the dispersions $\Delta_\varphi A$ and $\Delta_\varphi B$ are not at all associated with the experimental errors. There is

https://doi.org/10.1017/CBO9780511755361.004

nothing which *a priori* prevents $\langle A \rangle_\varphi$, for example, from being measured with an accuracy better than $\Delta_\varphi A$.

3. The position and momentum operators X and P (in one dimension) obey the commutation relation

$$[X, P] = \mathrm{i}\hbar I,$$

where $\hbar$ is the Planck constant, $\hbar = 1.054 \times 10^{-34}\,\mathrm{J\,s}$. Show that this commutation relation cannot be satisfied by operators acting in a Hilbert space of finite dimension. Hint: study the trace of this equation. From question **2** derive the Heisenberg inequality

$$\Delta X \Delta P \geq \frac{1}{2}\hbar.$$

2.7 Further reading

For additional information about light polarization and photons, the reader can consult Le Bellac (2006), Chapter 3, Lévy-Leblond and Balibar (1990), Chapter 4, and Hey and Walters (2003), Chapter 8. For the general principles of quantum mechanics see, for example, Nielsen and Chuang (2000), Chapter 2, which contains an elegant proof of the spectral decomposition theorem (Section 2.4). Some examples of determining a trajectory without perturbation in the Young slit experiment are given by Englert *et al.* (1991) and Dürr *et al.* (1998). A recent review on quantum cryptography containing numerous references to earlier work is that of Gisin *et al.* (2002). Popular accounts of quantum cryptography can be found in Bennett *et al.* (1992) and in Johnson (2003), Chapter 9. A very readable book on cryptography is Singh (2000).

https://doi.org/10.1017/CBO9780511755361.004