Chapter 1

Introducing Quantum Physics with Polarization

1.1 Polarization of a Light Beam

1.1.1 *Definition and basic measurement*

Light consists of electric and magnetic fields that can oscillate in any direction perpendicular to the direction of propagation. The polarization describes the direction of the electric field's oscillation. To measure this, we can use a polarizer, which is a material with a preferential axis due to its molecular or crystalline structure. This preferential axis allows the polarizer to act as a filter, transmitting only light polarized in certain directions.

We can liken polarized light being transmitted through a polarizer to metal bars with different orientations passing through a narrow door, which only allows vertically-oriented metal bars to pass through. Similarly, only light with certain polarization can be transmitted through a polarizer.

This analogy has limitations however, because only metal bars that are parallel to the door can pass through it, while those that are tilted from the vertical cannot. In comparison, polarization can be described as vectors, and even if the electric field is oriented at an angle to the polarizer's axis, the component parallel to the axis can still be transmitted, while the perpendicular component is reflected or dissipated as heat.

Imagine a polarizer with its axis oriented in the vertical direction. You send light through it, with an intensity I and polarization at an angle α to the vertical. Let the transmitted component be I_T, and the reflected component be I_R. I_T is oriented parallel to the vertical axis of the polarizer, whereas I_R is horizontal, perpendicular to the axis.

As intensity is a measure of energy, the law of conservation of energy give us the relation

$$I = I_T + I_R \,. \tag{1.1}$$

Next, imagine doubling the intensity I of the light that enters the polarizer. Since the initial intensity I is arbitrary, if it is doubled, the ratio of the transmitted and reflected intensities should still be the same. Hence I_T and I_R are also doubled. By intuition, the new relation should then be

$$2I = 2I_T + 2I_R\,. \tag{1.2}$$

We observe that $I_T \propto I$, and similarly $I_R \propto I$. Thus we obtain the relations

$$\begin{aligned} I_T &= k_1 I\,, \\ I_R &= k_2 I\,, \end{aligned} \tag{1.3}$$

where k_1, k_2 are constants. Substituting Equation (1.3) into Equation (1.1), we can deduce that

$$k_1 + k_2 = 1\,. \tag{1.4}$$

Two positive numbers such that $k_1 + k_2 = 1$ can always be written as $k_1 = \cos^2\alpha$ and $k_2 = \sin^2\alpha$, so finally

$$\begin{aligned} I_T &= I\cos^2\alpha\,, \\ I_R &= I\sin^2\alpha\,. \end{aligned} \tag{1.5}$$

This result is also known as Malus' Law in classical electromagnetic theory.

1.1.2 *Series of polarizers*

Let us consider a series of two polarizers, both with horizontal polarization axes (Figure 1.1). If the intensity transmitted by the first is I_{T1}, what is the intensity I_{T2} transmitted by the second? The answer is, of course, $I_{T2} = I_{T1}$, assuming that there is no energy loss when the light passes through the second polarizer.

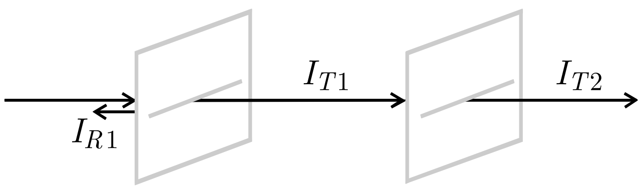

Fig. 1.1 Two polarizers with same polarization axes.

However, what happens if the second polarizer is tilted such that its polarization axis is now vertical (Figure 1.2)? The intensity I_{T2} now becomes zero.

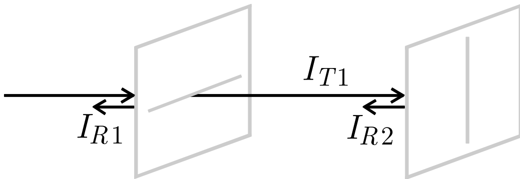

Fig. 1.2 Two polarizers with perpendicular polarization axes.

What happens next if a third polarizer is inserted between the two (Figure 1.3)? We expect that the intensity transmitted can decrease or remain the same if we add more polarizers, because more light would be filtered out, thus there should never be an increase. The curious thing is that, if this middle polarizer is oriented at a different angle from the other two, some light would pass through: the intensity actually increases! For example, if the middle polarizer has an axis oriented $\frac{\pi}{4}$ from the horizontal, the intensity transmitted by it would be $\frac{1}{2}I_{T1}$, and the intensity transmitted by the third polarizer would be half that from the second, that is, $\frac{1}{4}I_{T1}$.

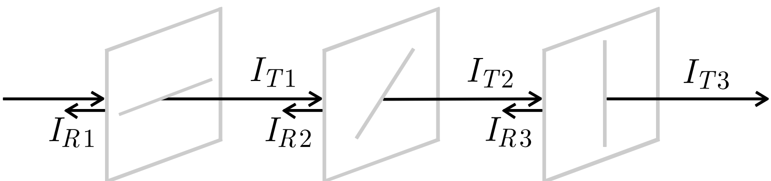

Fig. 1.3 Three polarizers each with polarization axis oriented $\frac{\pi}{4}$ with respect to the previous.

Hence polarizers are not merely filters that block a certain amount of light. This is a prediction of the laws and equations derived earlier, and it can easily be checked in a laboratory. The fact that this phenomenon has a surprising character when first encountered shows that our intuition sometimes fails, and we have to rely on observations and confirmed laws.

Now consider a case where we have the second and third polarizers oriented at angles α and β respectively from the first. How does the intensity transmitted by the third polarizer depend on α and β? From Equation (1.5), we know that the second polarizer will transmit a fraction $\cos^2\alpha$ of the intensity I_{T1} transmitted by the first. Since the third polarizer is oriented at $\beta - \alpha$ from the second, it would transmit a fraction $\cos^2(\beta - \alpha)$ of I_{T2}. Here, the important point to note is that the intensities depend only on the *difference* between the angles, and that the absolute orientation of the polarizer does not matter.

Exercise 1.1. A beam of horizontally-polarized light of intensity I impinges on a setup consisting of N polarizers, where N is very large. The first polarizer is oriented at an angle $\epsilon = \frac{\pi}{2N}$ to the horizontal; the next one is at an angle ϵ from the previous, and so on until the final one, which is exactly vertical. What is the intensity of light at the output of the last polarizer? What is its polarization? Note: neglect multiple reflections of the reflected beams and study only the transmitted beam.

1.1.3 *Polarization in vector notation*

In mechanics, it is very convenient to introduce a set of axes $(\hat{x}, \hat{y}, \hat{z})$ that will serve as coordinates; as well-known, the choice of such directions is completely arbitrary. Similarly, when studying polarization, it is customary to identify two orthogonal polarization directions as "horizontal" (H) and "vertical" (V). Any direction can be chosen for H; then, V will be the direction of polarization that is fully reflected by a polarizer that fully transmits H.

As usual, a direction can be described by a *vector*. Let us label the unit vectors along the vertical and horizontal directions as $\widehat{e}_V$ and $\widehat{e}_H$ respectively. Note that polarization is defined by the line along which the electric field oscillates, not by the direction along which the field points at a given time. Therefore, $-\widehat{e}_H$ describes the same polarization as $\widehat{e}_H$, and similarly for $\widehat{e}_V$.

With this basis of two vectors, one can describe any possible polarization direction: if the electric field is oscillating in a direction that makes an angle α with H, its polarization is[1]

$$\widehat{e}_\alpha = \cos\alpha\,\widehat{e}_H + \sin\alpha\,\widehat{e}_V \tag{1.6}$$

[1] Note that for simplicity we shall restrict our discussion to linear polarization throughout the main text; circular and elliptic polarization can be treated in a similar way by allowing complex numbers.

The polarization orthogonal to $\widehat{e}_\alpha$, i.e. the one that is fully reflected by a polarizer that fully transmits $\widehat{e}_\alpha$, is

$$\widehat{e}_{\alpha+\frac{\pi}{2}} \equiv \widehat{e}_{\alpha^\perp} = -\sin\alpha\,\widehat{e}_H + \cos\alpha\,\widehat{e}_V\,. \tag{1.7}$$

The remark we made on the previous page about the sign, of course, applies to all polarization vectors: $-\widehat{e}_\alpha$ describes the same polarization as $\widehat{e}_\alpha$. Note however that relative signs do matter: $\widehat{e}_\alpha = \cos\alpha\,\widehat{e}_H + \sin\alpha\,\widehat{e}_V$ describes a different polarization as $\cos\alpha\,\widehat{e}_H - \sin\alpha\,\widehat{e}_V \equiv \widehat{e}_{-\alpha}$.

1.1.4 *Polarizing beam-splitters as measurement devices*

We presented below the basic measurement of polarization, using a polarizer. For what follows, it will be rather convenient to keep in mind another device that measures polarization, namely the *polarizing beam-splitter* (Figure 1.4). Like the polarizer, this device splits light into two beams according to its polarization: for instance, horizontal polarization is transmitted and vertical polarization is reflected. However, here the reflected beam is not back-propagating or absorbed, as is the case with polarizers: rather, it is deflected in a different direction. This makes it easy to monitor both beams.

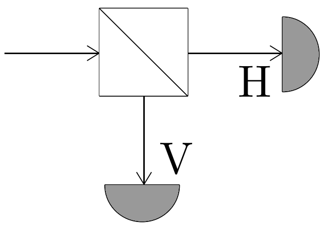

Fig. 1.4 Polarizing beam-splitter: scheme and implementation of a general measurement.

Note that one cannot rotate the device itself to measure along an angle α. In order to measure polarization along any arbitrary axis, polarizing beam-splitters must be complemented with *polarization rotators*. These are basically slabs of suitable materials, that can be chosen in order to perform the rotation

$$\widehat{e}_\alpha \longrightarrow \widehat{e}_H\,,$$
$$\widehat{e}_{\alpha^\perp} \longrightarrow \widehat{e}_V\,.$$

The measurement setup is now easy to understand: before the polarizing beam-splitter, one inserts the suitable rotator. If the output light is transmitted, then it was polarized along $\widehat{e}_H$ after the rotator, which means that it was polarized along $\widehat{e}_\alpha$ at the input. If on the contrary the output light is reflected, then it was polarized along $\widehat{e}_V$ after the rotator, which means that it was polarized along $\widehat{e}_{\alpha^\perp}$ at the input.

1.1.4.1 *Reconstructing polarization: tomography*

We conclude this section with an important remark on measurement. We have described measurement schemes that only allow one polarization to be distinguished from the orthogonal one. One might ask if there exist measurements that would give more detailed information: for instance, one that would discriminate perfectly between three or more possible polarization directions.

On a single beam of light, this is not possible; the measurements that we described are optimal. However, if the beam of light is intense, a much more refined measurement of polarization can be made: one just has to first split the beam into several beams without affecting the polarization. This can be done with half-transparent mirrors. On each of the beams, polarization can be measured along a different direction. In particular, if the total intensity is known, two measurements are sufficient to completely determine a linear polarization[2], as the following exercise shows. Such a measurement, from which the polarization direction can be fully reconstructed, is called *tomography*.

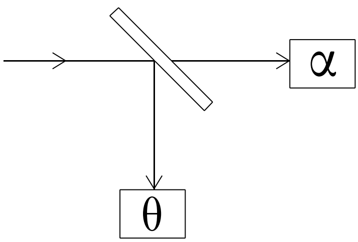

Fig. 1.5 Tomography of linear polarization.

[2] If the polarization is not guaranteed to be linear, three measurements would be needed.

Exercise 1.2. A beam of intensity $2I$ is split into two beams of intensity I each. On one of these beams, polarization is measured in the horizontal-vertical direction: the results are $I_H = I\cos^2\theta$, $I_V = I\sin^2\theta$. From these measurements, one can infer that the polarization is either $\widehat{e}_\theta = \cos\theta\,\widehat{e}_H + \sin\theta\,\widehat{e}_V$ or $\widehat{e}_{-\theta} = \cos\theta\,\widehat{e}_H - \sin\theta\,\widehat{e}_V$. Show that these two alternatives can be discriminated by measuring the polarization of the other beam in any other direction. Hint: suppose that the polarization is $\widehat{e}_\theta$: what are I_α and $I_{\alpha^\perp}$, with $\widehat{e}_\alpha$, $\widehat{e}_{\alpha^\perp}$ defined as in Equations (1.6), (1.7)? Compare these with the results obtained if the polarization is $\widehat{e}_{-\theta}$.

Using tomography, it seems that one can always learn the polarization direction perfectly. In classical physics, this is indeed the case. But the tomography process assumes that a light beam can always be split: what happens then, if the beam of light consists of the smallest unit, called a "photon"? This question marks the transition to quantum physics.

1.2 Polarization of One Photon

1.2.1 *Describing one photon*

During the development of modern physics, Einstein postulated that light is actually comprised of particles called photons, instead of the classical wave theory, to explain the photoelectric effect. The development of quantum physics clarified the notion of the photon. For the sake of this text, we can assume that light is "made" of photons.

In a polarized beam, each photon must have the same polarization. Indeed, if a polarizer is suitably oriented, the whole beam can be transmitted, that is each photon is transmitted.

Now, when dealing with a single photon, a different notation is used than when dealing with the whole beam. Namely, for a beam of polarization $\cos\alpha\,\widehat{e_H} + \sin\alpha\,\widehat{e_V}$, the state of polarization of each photon is written as $|\alpha\rangle = \cos\alpha|H\rangle + \sin\alpha|V\rangle$. This notation is called the *Dirac notation*, and we shall carefully study it now as we will be using it throughout the rest of the text.

1.2.2 *Computations using Dirac notation*

As mathematical objects, "$|.\rangle$" are vectors with very similar properties as the $\widehat{e}$ used for the field.

In particular, let us write

$$
\begin{aligned}
|H\rangle &\equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\
|V\rangle &\equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}.
\end{aligned} \tag{1.8}
$$

Then

$$|\alpha\rangle = \cos\alpha|H\rangle + \sin\alpha|V\rangle \equiv \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}. \tag{1.9}$$

The coefficient $\cos\alpha$ is, as in normal vector algebra, the *scalar product* of $|\alpha\rangle$ and $|H\rangle$. We write this scalar product as

$$\langle\alpha|H\rangle = \cos\alpha\,. \tag{1.10}$$

Similarly,

$$\langle\alpha|V\rangle = \sin\alpha\,. \tag{1.11}$$

In particular, by construction

$$\langle H|H\rangle = 1\,,\ \langle V|V\rangle = 1\,,\ \langle H|V\rangle = 0\,, \tag{1.12}$$

that is, $\{|H\rangle, |V\rangle\}$ is a basis.

From what we have discussed earlier, we know that for a beam the quantity $\cos^2\alpha$ represents the fraction of the beam intensity that is transmitted by a filter. But what does such a quantity mean at the level of single photons? A photon is an indivisible particle: when arriving at the filter, it can either be transmitted or reflected. Given this, $\cos^2\alpha$ represents the *probability* that the photon is transmitted. Thus we can state that the probability of a photon passing through horizontal and vertical polarizers, given an initial polarization angle α, is

$$
\begin{aligned}
P(\text{finding } H \text{ given } \alpha) = P(H|\alpha) = \cos^2\alpha = |\langle H|\alpha\rangle|^2\,, \\
P(\text{finding } V \text{ given } \alpha) = P(V|\alpha) = \sin^2\alpha = |\langle V|\alpha\rangle|^2\,.
\end{aligned}
$$

Indeed these two numbers sum up to 1 as they should for probabilities. The rule is valid for two arbitrary polarization states $|\psi_1\rangle$ and $|\psi_2\rangle$:

$$P(\text{finding } \psi_1 \text{ given } \psi_2) = |\langle\psi_1|\psi_2\rangle|^2. \tag{1.13}$$

This last equation is called *Born's rule for probabilities.*

It is important here to note that $|\alpha\rangle = \cos\alpha|H\rangle + \sin\alpha|V\rangle$ does not mean that some photons are polarized as $|H\rangle$ and some as $|V\rangle$. Rather, it describes a new polarization state whereby all the photons are polarized as $|\alpha\rangle$, but when they are measured in the H/V basis, they have probabilities $\cos^2\alpha$ and $\sin^2\alpha$ of being transmitted.

The appearance of the notion of probability in physics requires a careful discussion. Before that, to consolidate our understanding, we propose two exercises.

Exercise 1.3.

(1) Prove that $\{|\alpha\rangle, |\alpha^\perp\rangle\}$ is a basis for all α, where

$$\begin{aligned} |\alpha\rangle &= \cos\alpha|H\rangle + \sin\alpha|V\rangle\,, \\ |\alpha^\perp\rangle &= \sin\alpha|H\rangle - \cos\alpha|V\rangle\,. \end{aligned} \tag{1.14}$$

(2) Let $|\beta\rangle = \cos\beta|H\rangle + \sin\beta|V\rangle$. Compute the probabilities $P(\alpha|\beta)$ and $P(\alpha^\perp|\beta)$.

As discussed earlier, the probabilities depend only on the difference between the angles α and β, and not on their individual values. In general, when light polarized at an angle α to the horizontal passes through a polarizer with an axis oriented β to the horizontal, the parameter for determining the probability of transmission is $\alpha - \beta$, and

$$\begin{aligned} P(T) &= \cos^2(\alpha - \beta)\,, \\ P(R) &= \sin^2(\alpha - \beta)\,. \end{aligned} \tag{1.15}$$

1.2.3 *The meaning of probabilities*

Let us discuss in greater detail the meaning of probabilities in measurements on single photons. Consider a few photons with polarization $|\alpha\rangle = \cos\alpha|H\rangle + \sin\alpha|V\rangle$, moving towards a polarizer oriented in the horizontal direction. Let us first predict the results of this experiment. Based on the earlier analysis, we can determine the transmitted and reflected intensities to be

$$\begin{aligned} P(T) &= \cos^2\alpha\,, \\ P(R) &= \sin^2\alpha\,. \end{aligned} \tag{1.16}$$

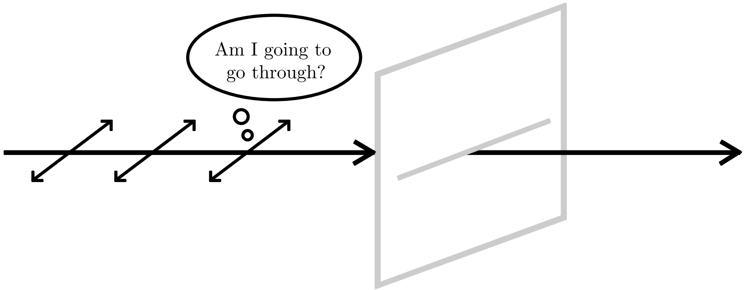

Fig. 1.6 Dilemma of a photon.

In other words, we can predict the statistical probability of the behaviour of a large number of photons, or a beam of light.

Now, let us consider each individual photon. Each photon is polarized along $|\alpha\rangle$ and is identical to each other. When one photon encounters the polarizer, what will happen to it? Will it pass through or will it be reflected? What about the next photon?

The answer is: we do not know. Although we know the statistical behaviour of many photons, we have no way of predicting the behaviour of each individual photon. Unsatisfied with not knowing the answer, we then ask, could it be that there is some hidden mechanism within each photon that would determine whether or not it would pass through?

The answer is, again, no. It has been experimentally proven that there is no such mechanism. In fact, we cannot predict the behaviour of each individual photon because the information is really not even there in the first place!

Our next question is, why can't this be predicted? Does this mean that nature is *random*? Does God play dice?

It may be surprising, but the answer is yes. The *randomness is intrinsic*. To make this point clearer, we can compare this situation to some well-known instances of randomness in our daily lives, such as tossing a coin, casting a die or weather patterns. In these cases, it may be hard or even impossible in practice to predict the results; however, this is not intrinsic randomness, since the relevant information is actually present.

The randomness in quantum phenomena is very different. We can measure the speed, polarization or any other parameters of a photon, but there is no way to tell whether or not it will pass through the polarizer. The information is really not there. This is what is meant by intrinsic randomness.

1.3 Describing Two Photons

Previously, in the case of a single photon, we were able to relate the statistical result of their intrinsically random behavior to the intensity of a light beam transmitted through a polarizer. However, two photons can exhibit purely quantum properties with no classical analogues.

1.3.1 *Classical composite systems*

In order to appreciate what follows, we consider first *composite systems* in classical physics. A physical system is composite if it can be seen as consisting of two or more subsystems. For example, the Earth and Moon. Neglecting its size, the physical properties of the Earth can be reduced to its position and speed $(\vec{x_E}, \vec{v_E})$; similarly for the Moon $(\vec{x_M}, \vec{v_M})$. Then the properties of the composite system can be derived from the set of vectors $(\vec{x_E}, \vec{v_E}, \vec{x_M}, \vec{v_M})$. The fact that the Earth and Moon interact means that, in time, the evolution of each parameter depends on the values of all the others. However, at any moment, the Earth has a well-defined position and speed; similarly for the Moon. This may seem obvious, and indeed it is in classical physics. But in quantum physics, this is not always the case: the properties of each subsystem may not be well-defined.

1.3.2 *Four states of two photons*

Consider now a composite system of two photons. Both photons could be horizontally polarized, or both vertically, or the first one horizontally and the second vertically, or the opposite. These situations are written as $|H\rangle \otimes |H\rangle$, $|V\rangle \otimes |V\rangle$, $|H\rangle \otimes |V\rangle$, and $|V\rangle \otimes |H\rangle$ respectively. The symbol $\otimes$ is a multiplication between vectors known as the tensor product defined as follows: let

$$|H\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ |V\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{1.17}$$

then

$$|H\rangle \otimes |H\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 0\begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{1.18}$$

and by using the same technique,

$$|H\rangle \otimes |V\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |V\rangle \otimes |H\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |V\rangle \otimes |V\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{1.19}$$

Take note that this multiplication is not commutative, i.e $|H\rangle \otimes |V\rangle \neq |V\rangle \otimes |H\rangle$. This is obvious if you recall the physical meaning of this notation: "the first photon is H and the second is V" is definitely not the same as "the first photon is V and the second is H". Apart from this, the tensor product has all the usual properties of multiplication.

For simplicity, we shall write $|H\rangle \otimes |H\rangle$ as $|H, H\rangle$ or $|HH\rangle$, a convention that we will be using throughout this book.

1.3.3 *All the states of two photons*

Previously, we saw that the general expression for the linear polarization state of a single photon is $|\alpha\rangle = \cos\alpha|H\rangle + \sin\alpha|V\rangle$, and any α defines a valid state. Similarly, in the case of two photons, the most general state takes the form

$$|\psi\rangle \;=\; a|HH\rangle + b|HV\rangle + c|VH\rangle + d|VV\rangle\,. \tag{1.20}$$

To keep the interpretation of $|a|^2$, $|b|^2$, $|c|^2$ and $|d|^2$ as probabilities, we also impose the *normalization condition* $|a|^2 + |b|^2 + |c|^2 + |d|^2 = 1$. Conversely, any choice of $\{a, b, c, d\}$ satisfying the normalization condition represents a valid state.

Let us now study some examples. We start with

$$c|HH\rangle + s|HV\rangle = |H\rangle(c|H\rangle + s|V\rangle)\,, \tag{1.21}$$

where $c = \cos\alpha$ and $s = \sin\alpha$. This state represents the situation in which the first photon has a polarization of $|H\rangle$ while the second photon has a polarization of $|\alpha\rangle = c|H\rangle + s|V\rangle$.

Let us take a look at another state: $\frac{1}{\sqrt{2}}(|HH\rangle + |VV\rangle)$. Intuitively, this should also correspond to each photon having a well-defined polarization:

$$\frac{1}{\sqrt{2}}(|HH\rangle + |VV\rangle) \stackrel{?}{=} (c_\alpha|H\rangle + s_\alpha|V\rangle) \otimes (c_\beta|H\rangle + s_\beta|V\rangle)\,, \tag{1.22}$$

where $c_\alpha = \cos\alpha$ etc. Expanding the product, we get

$$\begin{aligned}\frac{1}{\sqrt{2}}|HH\rangle + \frac{1}{\sqrt{2}}|VV\rangle \stackrel{?}{=} c_\alpha c_\beta|HH\rangle + c_\alpha s_\beta|HV\rangle \\ + s_\alpha c_\beta|VH\rangle + s_\alpha s_\beta|VV\rangle\,.\end{aligned} \tag{1.23}$$

To satisfy this equality, we must have $c_\alpha c_\beta = s_\alpha s_\beta = \frac{1}{\sqrt{2}}$, but $c_\alpha s_\beta = s_\alpha c_\beta = 0$: the two conditions are manifestly contradictory. Hence, there is no way to write this state as the product of two independent states, one for each subsystem.

However, $\frac{1}{\sqrt{2}}(|HH\rangle + |VV\rangle)$ still describes a valid physical state, as we have mentioned earlier. We have to admit that there are some states for which it is not possible to describe the two photons separately. Such states are called *entangled*. It turns out that the conditions for composite states to be written as the product of two separate states are very constraining, and that vectors chosen at random are not likely to satisfy this. Thus most composite states are entangled.

Exercise 1.4. Consider the following two-photon states:

$$\begin{aligned}|\psi_1\rangle &= \frac{1}{2}(|HH\rangle + |HV\rangle + |VH\rangle + |VV\rangle)\,,\\ |\psi_2\rangle &= \frac{1}{2}(|HH\rangle + |HV\rangle + |VH\rangle - |VV\rangle)\,,\\ |\psi_3\rangle &= \frac{1}{2}|HH\rangle + \frac{\sqrt{3}}{2}(|VH\rangle + |VV\rangle)\,,\\ |\psi_4\rangle &= \cos\theta|HH\rangle + \sin\theta|VV\rangle\,.\end{aligned}$$

Verify that all the states are correctly normalized. Which ones are entangled? Write those that are *not* entangled as an explicit product of single-photon states.

1.3.4 *The meaning of entanglement*

What does entanglement mean? Let us stress explicitly that we are studying states, i.e. the description of the system, not its dynamics. The Earth and the Moon mentioned earlier influence the *motion* of each other, but we stressed (it seemed obvious) that at each time one can assign a state to the Earth and a state to the Moon. On the contrary, an entangled state describes a situation in which, at a given time and without any mention of evolution, two photons cannot be described separately. How should we understand this?

It is very hard, if not impossible, to give an intuitive meaning to entanglement since it is not observed in everyday life. However, physicists have performed tests on entangled photons and observed their properties. Some of these tests would be described in detail in the following chapters. We conclude with an intriguing exercise:

Exercise 1.5. Consider $|\alpha\rangle$ and $|\alpha^\perp\rangle$ as defined in Equation (1.14). Check that

$$\frac{1}{\sqrt{2}}\big(|\alpha\rangle|\alpha\rangle + |\alpha^\perp\rangle|\alpha^\perp\rangle\big) = \frac{1}{\sqrt{2}}\big(|H\rangle|H\rangle + |V\rangle|V\rangle\big)\,. \tag{1.24}$$

What happens if both photons are measured in the same basis?

Notice that the polarization of the two photons are always the same, for all measurements. We say that their polarization are perfectly correlated. However, this correlation is not caused by each photon having the same well-defined state, for example, one photon is $|H\rangle$, the other is also $|H\rangle$, thus both always have the same polarization. Recall that for entangled photons, each photon cannot be described separately. This may be counter-intuitive, but even though both photons do not have well-defined properties, their properties are correlated such that their relative states are well-defined. Using the same example, this means that although we cannot say that either photon is $|H\rangle$, if one is measured to be $|H\rangle$, the other must definitely be $|H\rangle$ as well. Thus we see the presence of correlations in the properties of entangled photons.

1.4 Transformations of States

We have discussed how to measure states and predict the results of the measurements. Another important concept is the transformation of states, and the properties that such transformations obey.

First, we have to stress that the measurement process discussed thus far is an optimal one. In other words, there is no way in which two states can be distinguished better than by measuring them using polarizers or polarizing beam splitters. This is important because otherwise, Born's rule of probabilities would not be accurate as there are more optimal measurements that will produce different probabilities.

Secondly, it is an assumption of modern physics that transformations between macroscopic states are *reversible*. This means that if a state $|\phi\rangle$ can be transformed to $|\theta\rangle$, $|\theta\rangle$ can also be transformed to $|\phi\rangle$. An example of this is the rotation of a vector: it can be reversed by a rotation in the opposite direction. It turns out that all reversible transformations on a polarization vector are indeed related to rotations.

A theorem due to Wigner proves that in quantum physics, all transformations must be *linear* and have the property of *unitarity*:

- A transformation T is linear if, for all pairs of vectors $\{|\psi_1\rangle, |\psi_2\rangle\}$,
$$T(|\psi_1\rangle + |\psi_2\rangle) = T(|\psi_1\rangle) + T(|\psi_2\rangle)\,. \tag{1.25}$$
- A transformation U is unitary if it preserves the scalar product of any two vectors:
$$\langle\phi_1|\phi_2\rangle = \langle U(\phi_1)|U(\phi_2)\rangle\,. \tag{1.26}$$

The proof of linearity is the most difficult part of Wigner's theorem, here we sketch the reason behind the necessity of preserving the scalar product.

According to Born's rule, two states are perfectly distinguishable if the scalar product $\chi = 0$, and are the same if $\chi = 1$, indicating that χ is a measure of the "distinguishability" of the two states in an optimal measurement.

Now, let us analyze what would happen if χ is not preserved. If we suppose that there is a transformation such that χ decreases, this means that the states become more distinguishable, which contradicts the assumption that the measurement is optimal. Indeed, before performing the measurement, we could apply the transformation. If instead χ increases, the assumption of reversibility implies that there exists a reversed transformation in which χ decreases. Thus the only remaining option is that χ does not change during a reversible evolution.

1.5 Summary

The polarization of light describes the direction of the electric field's oscillation, and is written using vectors. Polarizers or polarizing beam-splitters are used to determine the polarization, by measuring the transmitted and reflected light intensities. For single photons, their polarization states are written using the Dirac notation, and Born's rule determines the probability of transmission or reflection through a measuring device. The behavior of single photons is intrinsically random; only the statistical behavior of a large number of photons can be predicted. Two photons can form entangled states in which each photon does not have a well-defined state, but are correlated such that their relative states are well-defined. Reversible transformations are also possible for states; these must be linear and unitary.

1.6 The Broader View

In this section, we present some of the concepts mentioned in this chapter in greater depth. These include vectors, probability and the degree of freedom.

1.6.1 *Vectors*

Vectors are a basic mathematical concept used in quantum theory. General vectors have similar properties as the geometric vectors encountered in elementary mathematics courses. The main difference is the notation: instead of the commonly used $\vec{v}$, in quantum mechanics we use the Dirac notation, $|v\rangle$. We also write the scalar product as $\langle v_1, v_2\rangle$ or $\langle v_1|v_2\rangle$ instead of $\vec{v_1}\cdot\vec{v_2}$.

Vectors can be expressed as a sum of component vectors, which is usually done using unit vectors that form a basis. For example,

$$\begin{pmatrix} a \\ b \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \end{pmatrix}\,. \tag{1.27}$$

In the Dirac notation, $\begin{pmatrix} a \\ b \end{pmatrix}$ is written as $|v\rangle$, and $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ as $|e_1\rangle$ and $|e_2\rangle$ respectively. Then Equation (1.27) can be written as

$$|v\rangle = a|e_1\rangle + b|e_2\rangle\,. \tag{1.28}$$

Notice that a is actually the scalar product of $|v\rangle$ and $|e_1\rangle$, or $a = \langle e_1|v\rangle$. Likewise, $b = \langle e_2|v\rangle$. $|e_1\rangle$ and $|e_2\rangle$ also form an orthonormal basis, which means that these conditions are satisfied: $\langle e_1|e_1\rangle = 1$, $\langle e_2|e_2\rangle = 1$ and $\langle e_1|e_2\rangle = 0$. Geometrically, $|e_1\rangle$ and $|e_2\rangle$ are orthogonal to each other and normalized to one.

1.6.2 *Probability*

Many events in our daily lives are random, e.g. the results of tossing a coin or casting a dice. That is, their behaviour is unpredictable in the short term. However, these events may have a regular and predictable pattern in the long term. In the case of tossing a coin, we find that after many tries, half of the tosses produces heads and the other half produces tails. The *probability* is the proportion of times an event occurs over a large number of repetitions. Mathematically, we define the probability of an event A as

$$P(A) = \frac{n(A)}{n(S)}\,, \tag{1.29}$$

where $n(A)$ is the number of times event A occurs and $n(S)$ is the total number of events.

The probability $P(A)$ of any event A satisfies $0 \leq P(A) \leq 1$. By definition, if $P(A) = 1$, then event A will definitely happen and if $P(A) = 0$, then event A will definitely not happen.

Two events A and B are independent if the probability of one event happening does not affect the probability of the other event happening. For example, if you toss a coin and the result is a tail, the result of this toss does not affect the result of the next toss. Hence each toss is independent. If A and B are independent, then $P(A \cap B) = P(A) \times P(B)$, where $A \cap B$ denotes that both events A and B occur.

1.6.3 *Degree of freedom*

You may have encountered the phrase "degree of freedom" in chemistry or physics courses, in which mono-atomic particles are described as having three degrees of freedom while diatomic particles have five. In this case, each degree of freedom is a type of motion that the particles can undergo. Here in this book, this term will be used more generally to denote a physical variable that describes the system under study.

Within a system, there are often many variables, but not all the variables are of interest. For example, if we have identified a car as our system, the variables would include its position, velocity, and also its colour, temperature, material etc. If our aim is to study the motion of the car, then the colour of the car would be of little significance and can be ignored. In contrast, the position and speed of the car are of great importance; we call them the *degrees of freedom* of the car.

Let us take another system as an example. Suppose the physical system is an electron and our aim is to examine its motion. Again, the speed and position of the electron are its degrees of freedom. Its motion can be described by the equation $m\ddot{\vec{x}} = -e\vec{E}(x)$, where $\vec{E}(x)$ is the electric field in the region, and e is the charge of the electron.

However, if instead we want to take into account how the electric field changes as the electron moves, our system would then be composed of both the electron and electric field. We would then have to express their relationship using Maxwell's equations, and the analysis becomes more complicated. Thus as the system under consideration changes, the method of analysis also changes.

Before we end this discussion, we illustrate an important difference between quantum and classical systems through set theory. We first consider a classical system: in the example of the car, we can describe its degrees of freedom using a velocity-position graph, as in Figure 1.7.

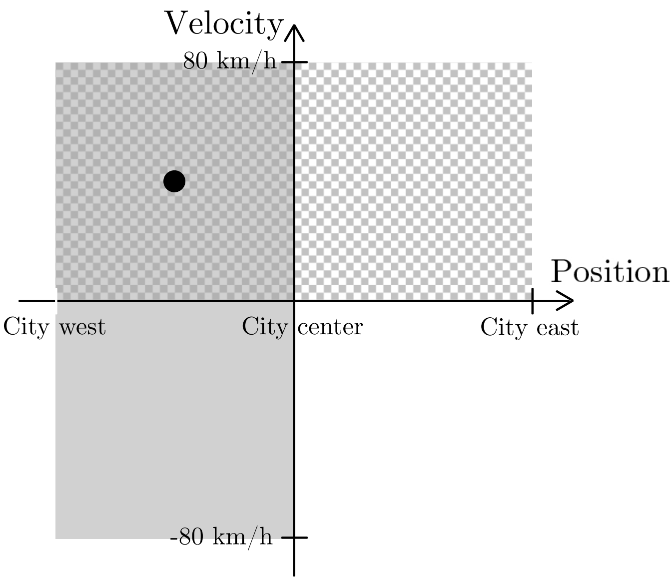

Fig. 1.7 Velocity-position graph.

The shaded region indicates all the cars that are between the west of the city and the city center, while the patterned region indicates all the cars that are traveling towards the east. By simple logic based on set theory, the overlapped region indicates the cars that are both between the west of the city and the city center, and are traveling towards the east. Thus if the car is at the position of the dot, we know that the car is between the west of the city and the city center, and is traveling at about 40 km/h towards the city center. We observe from this example that classical properties can be described using set theory.

To generalize this concept, suppose P_1 and P_2 are two physical properties, while S_1 and S_2 are sets that satisfy the physical properties P_1 and P_2 respectively. In mathematical notation,

$$\begin{aligned} P_1 &\to S_1\,, \\ P_2 &\to S_2\,. \end{aligned} \tag{1.30}$$

From what we have discussed thus far, we know that

$$\begin{aligned} 1.\, P_1 \wedge P_2 \ &\to\ S_1 \cap S_2\,, \\ 2.\, P_1 \vee P_2 \ &\to\ S_1 \cup S_2\,, \\ 3.\, P_1 \Rightarrow P_2 &\to S_1 \subset S_2\,. \end{aligned} \tag{1.31}$$

In words, the intersection of S_1 and S_2 satisfies the property P_1 and P_2, the union of S_1 and S_2 satisfies the property P_1 or P_2, and if S_1 is a subset of S_2, P_2 is satisfied if P_1 is satisfied.

However, this simple, intuitive relationship breaks down in quantum theory, because quantum systems do not have well-defined properties.

1.7 General References

Popular books on quantum physics:

- G.C. Ghirardi, *Sneaking a Look at God's Cards* (Princeton University Press, Princeton, 2003).
- V. Scarani, *Quantum Physics: A First Encounter* (Oxford University Press, Oxford, 2006).

General advanced textbooks:

- On classical optics: E. Hecht, *Optics* (Addison Wesley, San Francisco, 2002).
- A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, 1995).
- M. Le Bellac, *A Short Introduction to Quantum Information and Quantum Computation* (Cambridge University Press, Cambridge, 2006).

1.8 Solutions to the Exercises

Solution 1.1. To solve this problem, first note that the intensity transmitted by each polarizer is a fraction $\cos^2\epsilon$ of the previous. Thus to find the intensity transmitted by the last one, we can multiply all these factors together:

$$\cos^2\epsilon \times \cos^2\epsilon \times \ldots = (\cos^2\epsilon)^N, \text{ where } N = \frac{\pi}{2\epsilon}\,.$$

We can then take the limit of the term $(\cos^2\epsilon)^N = (\cos^2\epsilon)^{\frac{\pi}{2\epsilon}}$ as ϵ approaches zero, and we find that

$$\lim_{\epsilon\to 0}(\cos\epsilon)^{\frac{\pi}{\epsilon}} = 1\,.$$

Thus we conclude that the intensity transmitted by the last polarizer is equal to the original intensity I! But the direction of polarization has changed: the light entered horizontally polarized and exited vertically polarized. In effect, this setup rotated the polarization while maintaining the intensity. A cultural remark: this effect, especially in the context of quantum physics, is called the *Zeno effect* from the name of a Greek philosopher; we let you guess why.

Solution 1.2. Suppose that the polarization is $\widehat{e}_\theta$, we can find I_α and $I_{\alpha^\perp}$ with reference to Equation (1.5), in which the difference in angles is now $\theta - \alpha$.

$$\begin{aligned} I_\alpha &= I\cos^2(\theta-\alpha)\,, \\ I_{\alpha^\perp} &= I\sin^2(\theta-\alpha)\,. \end{aligned}$$

Similarly, if the polarization is $\widehat{e}_{-\theta}$, then I_α and $I_{\alpha^\perp}$ are given as follows:

$$\begin{aligned} I_\alpha &= I\cos^2(\theta+\alpha)\,, \\ I_{\alpha^\perp} &= I\sin^2(\theta+\alpha)\,. \end{aligned}$$

This difference in the resulting I_α and $I_{\alpha^\perp}$ can be measured, thus the two polarization directions can be distinguished.

Solution 1.3.

(1) A basis is defined by

$$\begin{aligned} \langle\alpha|\alpha\rangle &= \cos^2\alpha + \sin^2\alpha = 1\,, \\ \langle\alpha^\perp|\alpha^\perp\rangle &= \cos^2\alpha + \sin^2\alpha = 1\,, \\ \langle\alpha|\alpha^\perp\rangle &= \cos\alpha\sin\alpha - \cos\alpha\sin\alpha = 0\,. \end{aligned}$$

These conditions can indeed be verified, either by writing both $|\alpha\rangle$ and $|\alpha^\perp\rangle$ as in Equation (1.9), or directly by using Equation (1.12).

(2) Using Born's rule for probabilities,

$$P(\alpha|\beta) = |\langle\alpha|\beta\rangle|^2 = |\cos\alpha\cos\beta + \sin\alpha\sin\beta|^2 = \cos^2(\alpha-\beta)\,,$$
$$P(\alpha^\perp|\beta) = |\langle\alpha^\perp|\beta\rangle|^2 = |\sin\alpha\cos\beta - \cos\alpha\sin\beta|^2 = \sin^2(\alpha-\beta)\,.$$

Solution 1.4. $|\psi_1\rangle$ is *not* entangled: it can be written as $\frac{1}{\sqrt{2}}(|H\rangle + |V\rangle) \otimes \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle) = |\alpha = \frac{\pi}{4}\rangle \otimes |\beta = \frac{\pi}{4}\rangle$. All the other states are entangled (of course, $|\psi_4\rangle$ is not entangled if $\cos\theta = 0$ or $\sin\theta = 0$, but it is in all the other cases).

Solution 1.5.

$$\begin{aligned}
&\frac{1}{\sqrt{2}}\Big[|\alpha\rangle|\alpha\rangle + |\alpha^\perp\rangle|\alpha^\perp\rangle\Big] \\
&= \frac{1}{\sqrt{2}}\Big[(c|H\rangle + s|V\rangle)(c|H\rangle + s|V\rangle) + (s|H\rangle - c|V\rangle)(s|H\rangle - c|V\rangle)\Big] \\
&= \frac{1}{\sqrt{2}}\Big[(c^2+s^2)|HH\rangle + (cs-sc)|HV\rangle + (sc-cs)|VH\rangle + (s^2+c^2)|VV\rangle\Big] \\
&= \frac{1}{\sqrt{2}}\Big[|HH\rangle + |VV\rangle\Big].
\end{aligned}$$