Polarization State of a Photon

S. Turgut

September 25, 2003

In here we give a quantum mechanical description of the polarization state of a photon. In my opinion, the material explained here is easy to understand and for that reason it is suitable for an introduction to some of the concepts of quantum mechanics. First we will see the representation of a state as a vector, then we will discuss the effect of a measurement on a state vector. After that operators for physical observables are introduced.

The discussion in what follows is exactly same for any two-level system like the spin state of an electron or the states of a two-level atom. Only the physical meaning will be different. The main point of choosing the polarization state of a photon, instead of electron spin is that the latter is not familiar to the students. As a result, some of the experimental results has to be given as postulates. On the other hand, if somebody has played with a pair of polarizers before, he/she will understand most of the "postulates" in a natural way.

1 Polarization State of a Classical Electromagnetic Wave

Before seeing the description of the polarization state of a photon, we need to see the description of that of a classical wave. The main reason of doing this is a basic principle of quantum mechanics. The nature is basically working with quantum mechanics but everything macroscopic can be treated successfully with classical mechanics. For some reason quantum laws operating in microscopic domain give rise to classical macroscopic laws. This is the so-called *Correspondence Principle*. As a result, millions of photons in the same state (or almost the same state) can be successfully described as a classical electromagnetic wave. This enables us to "derive" some of the laws of quantum mechanics from classical considerations.

The waves from a typical source (like Sun) are unpolarized. But if those waves pass from a polarizer, the resulting wave will be linearly polarized. It is possible to prepare a wave in other polarization states as well: Linearly, circularly or elliptically polarized waves can be obtained with suitable apparatus. We will assume that we are working with such waves. Description of unpolarized (or partially polarized) waves is another problem.

1.1 Matrix Representation of Polarization States

A linearly polarized plane wave will have the associated fields

$$\vec{\mathbf{E}}(\vec{\mathbf{r}},t) = \vec{\mathbf{E}}_0 \cos(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t + \phi) \quad , \tag{1}$$

$$\vec{\mathbf{B}}(\vec{\mathbf{r}},t) = \vec{\mathbf{B}}_0 \cos(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t + \phi) \quad , \tag{2}$$

Here $\vec{\mathbf{k}}$ is the wavevector of the wave, ϕ is some phase angle, ω is the angular frequency, and $\vec{\mathbf{E}}_0$ and $\vec{\mathbf{B}}_0$ are some constant vectors (amplitudes) describing the polarization state of the wave.

In writing this expression we have already made an assumption. We assumed that the wave is a plane wave rather than a real wave which is localized in space. A real wave occupies a finite volume, and as a result of this, it has a range of different $\vec{\mathbf{k}}$ values. But since we are interested in the polarization state of the wave, we can ignore this part related to the "positional state".

We know that the Maxwell's equations imply that $\vec{\mathbf{k}}$, $\vec{\mathbf{E}}_0$ and $\vec{\mathbf{B}}_0$ are all perpendicular to each other. Suppose that the wave is sent along a chosen direction which we take as our z-axis. That means that the electric and magnetic fields are on the $x-y$ plane. Moreover, knowledge of the electric field is enough to determine the magnitude and direction of the magnetic field (in Gaussian units $\vec{\mathbf{B}}_0 = \hat{\mathbf{k}} \times \vec{\mathbf{E}}_0$). As a result, it is enough to specify the electric field to determine the polarization state of the wave. If we choose a fixed x and y axes, the electric field amplitude can be expressed in terms of two constants as

$$\vec{\mathbf{E}}_0 = E_{0x}\hat{\mathbf{x}} + E_{0y}\hat{\mathbf{y}} \quad . \tag{3}$$

As a result two real numbers have to be given to express the fields of a linearly polarized wave whose propagation direction is known. And, once you find these two numbers, you would have determined the polarization state of the wave.

At this point, we will use an unfamiliar mathematical tool to denote the wave given in Eq. (1) as a 2×1 matrix (or a two-dimensional vector) as

$$W = \begin{bmatrix} E_{0x}e^{i\phi} \\ E_{0y}e^{i\phi} \end{bmatrix} \quad . \tag{4}$$

This matrix is called the *Jones vector* in the optics community. We will use this notation as a good jumping point for the representation of the *state vector* for photons. At this point just note that from the Jones vector in (4) we can recover the wave in (1). If a wave has the Jones vector

$$W = \begin{bmatrix} A \\ B \end{bmatrix}$$

where A and B are two complex numbers in the units of electric-field, then the electric field at any point $\vec{\mathbf{r}}$ and at any time t is given by

$$\vec{\mathbf{E}}(\vec{\mathbf{r}}, t) = \mathrm{Re}\left(Ae^{i\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - i\omega t}\right)\hat{\mathbf{x}} + \mathrm{Re}\left(Be^{i\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - i\omega t}\right)\hat{\mathbf{y}}$$

As a result, we can easily change the notations back and forth. Note that the inclusion of the phase ϕ in Eq. (4) is necessary if you want to recover the expression of the electric field fully.

With this expression we have already extended our notation to arbitrary polarization. For an arbitrary polarization, the x and y components of the electric field will have different phase factors. If the electric field is expressed as

$$\vec{\mathbf{E}}(\vec{\mathbf{r}}, t) = E_{0x}\cos(\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} - \omega t + \phi_1)\hat{\mathbf{x}} + E_{0y}\cos(\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} - \omega t + \phi_2)\hat{\mathbf{y}} \tag{5}$$

then the matrix representation of the wave will be

$$\begin{bmatrix} E_{0x}e^{i\phi_1} \\ E_{0y}e^{i\phi_2} \end{bmatrix} \quad . \tag{6}$$

1.2 Typical Polarization States

Let us then see the typical polarization states. The simplest are the ones which are linearly polarized along x and y axes. For the linear polarization along x axis (we will call this *horizontal* polarization) the matrix is

$$\begin{bmatrix} Ee^{i\phi} \\ 0 \end{bmatrix} = Ee^{i\phi}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

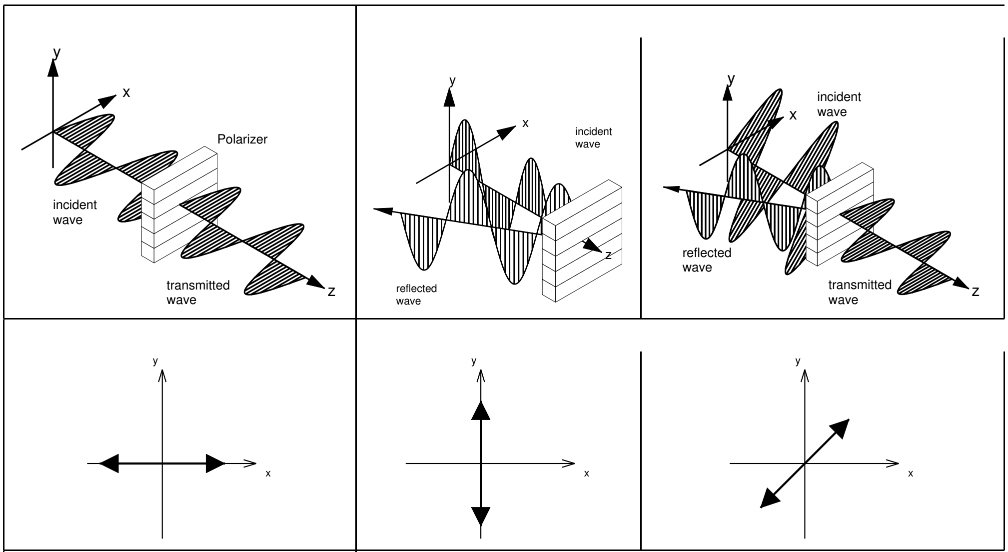
and for the polarization along y axis (we will call that *vertical* polarization) the matrix is

$$\begin{bmatrix} 0 \\ Ee^{i\phi} \end{bmatrix} = Ee^{i\phi}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad .$$

In both cases, E is a real number giving the maximum value of the electric field.

Next, let us express the matrix for linear polarization along an arbitrary axis $\hat{\mathbf{n}}$ where $\hat{\mathbf{n}}$ is a unit vector on the xy plane. If $\hat{\mathbf{n}}$ makes an angle θ with the x-axis (we assume the usual convention for the positive direction for θ) then

$$\hat{\mathbf{n}} = n_x\hat{\mathbf{x}} + n_y\hat{\mathbf{y}} = \cos\theta\hat{\mathbf{x}} + \sin\theta\hat{\mathbf{y}} \quad .$$



Horizontal polarization	Vertical Polarization	Arbitrary linear polarization

Table 1: Linearly polarized light incident on a polarizer. We assume that the perpendicularly polarized light is reflected, not absorbed.

If the electric field is given as

$$\vec{\mathbf{E}}(\vec{\mathbf{r}},t) = E\cos(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t + \phi)\hat{\mathbf{n}}$$

then the Jones vector is

$$\left[\begin{array}{c} En_x e^{i\phi} \\ En_x e^{i\phi} \end{array}\right] = Ee^{i\phi}\left[\begin{array}{c} n_x \\ n_y \end{array}\right] = Ee^{i\phi}\left[\begin{array}{c} \cos\theta \\ \sin\theta \end{array}\right]$$

Circular polarization is slightly more complicated. In this case, the x and y components of the wave will be phase shifted by 90^o or $\pi/2$ radians. When the electric field is observed at a definite position (say $\vec{\mathbf{r}} = 0$), it will be seen to be rotating as time passes without changing its magnitude. In the case of *right circularly polarized* wave, the electric field rotates clockwise. Alternatively, when you look at the fields at a fixed time (say $t = 0$), the tips of the electric field draw a right-handed spiral. The expression for the electric field is then

$$\vec{\mathbf{E}}(\vec{\mathbf{r}},t) = E\cos(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t + \phi)\hat{\mathbf{x}} + E\sin(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t + \phi)\hat{\mathbf{y}} \quad .$$

Note that the magnitude of the field does not change with time. Since $\sin\xi = \cos(\xi - \pi/2)$, we can rewrite this expression as

$$\vec{\mathbf{E}}(\vec{\mathbf{r}},t) = E\cos(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t + \phi)\hat{\mathbf{x}} + E\cos(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t + \phi - \frac{\pi}{2})\hat{\mathbf{y}} \quad .$$

As a result, the matrix for *right circularly polarized* wave is

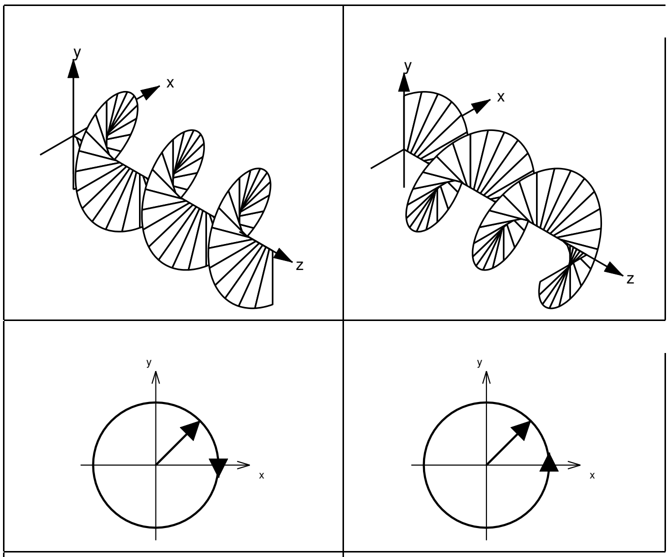
$$\left[\begin{array}{c} Ee^{i\phi} \\ Ee^{i(\phi-\frac{\pi}{2})} \end{array}\right] = Ee^{i\phi}\left[\begin{array}{c} 1 \\ -i \end{array}\right] \quad .$$

For the case of *left circularly polarized* wave, the tips of electric field vectors draw a left-handed spiral at constant time. The electric field is given by

$$\begin{aligned} \vec{\mathbf{E}}(\vec{\mathbf{r}},t) &= E\cos(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t + \phi)\hat{\mathbf{x}} - E\sin(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t + \phi)\hat{\mathbf{y}} \\ &= E\cos(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t + \phi)\hat{\mathbf{x}} + E\cos(\vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t + \phi + \frac{\pi}{2})\hat{\mathbf{y}} \end{aligned} \tag{7}$$

and the matrix representation is

$$\left[\begin{array}{c} Ee^{i\phi} \\ Ee^{i(\phi+\frac{\pi}{2})} \end{array}\right] = Ee^{i\phi}\left[\begin{array}{c} 1 \\ i \end{array}\right] \quad .$$



Right circular polarization | Left circular polarization

1.3 Superposition

When two coherent waves are passing over the same region, the fields of the waves are added up vectorially. There are practical ways of doing this for light beams. A beam might be split up into two by using a semi-reflecting glass plate, after the two beams pass from different regions, they can be joined by a similar glass plate. In here we are assuming that the two beams are moving along the same direction (their $\vec{\mathbf{k}}$ vectors are parallel) so that we can talk about a single resultant wave formed by superposition. In this case, the Jones vectors of the two waves are simply added up. This is a method of changing the polarization state of a wave (not the best way).

As an example, consider two horizontally polarized waves. Let

$$\vec{\mathbf{E}}_\ell(\vec{\mathbf{r}}, t) = E_\ell \cos(\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} - \omega t + \phi_\ell)\hat{\mathbf{x}}$$

for $\ell = 1, 2$. Then, if you define E and ϕ such that

$$E_1 e^{i\phi_1} + E_2 e^{i\phi_2} = E e^{i\phi}$$

you can show that

$$\vec{\mathbf{E}}_1(\vec{\mathbf{r}}, t) + \vec{\mathbf{E}}_2(\vec{\mathbf{r}}, t) = E \cos(\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} - \omega t + \phi) \quad .$$

In other words, if you add up the Jones vectors for the two waves, you will obtain the Jones vector for the superposed wave

$$\begin{bmatrix} E_1 e^{i\phi_1} \\ 0 \end{bmatrix} + \begin{bmatrix} E_2 e^{i\phi_2} \\ 0 \end{bmatrix} = \begin{bmatrix} E e^{i\phi} \\ 0 \end{bmatrix} \quad .$$

This can be generalized to waves with arbitrary polarization. To sum up, superposition corresponds to adding up the Jones vectors.

1.4 Effect of Polarizers

A polarizer is a solid material which is transparent to waves of one polarization and opaque to the "other" polarization. Below we are going to specify what do we mean by "other". For a linear polarizer, the two polarization directions are perpendicular to each other. As a result, a polarizer oriented along x-axis permits waves with horizontal polarization to pass through and completely absorbs the vertically polarized waves. Absorption is not an essential feature. We may also construct polarizers where the waves are either *transmitted* or *reflected* with no possibility of absorption. Opticians call such instruments Polarizing Beam Splitters (PBS). Below, I am going to assume that this is the case.

Suppose that a wave with arbitrary polarization is incident on a polarizer and it has the Jones vector

$$W = \begin{bmatrix} A \\ B \end{bmatrix}$$

where A and B are complex numbers. We can think of this wave as a superposition of two different waves with different polarization

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} A \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ B \end{bmatrix} \quad ,$$

and as a result analyze the behavior of two waves separately. In this case, only the first wave with amplitude A will be able to pass the polarizer. We are going to represent this by

$$\begin{bmatrix} A \\ B \end{bmatrix} \longrightarrow \begin{bmatrix} A \\ 0 \end{bmatrix} \qquad \text{transmitted wave.}$$

As a result, passing a polarizer corresponds to a *projection*. By a projection we mean that we eliminate all components of a vector and exactly copy the remaining components. For the transmitted wave, the projection is onto the x axis. In optics, the effect of the polarizer is sometimes represented by a matrix multiplication.

$$\begin{bmatrix} A \\ B \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} A \\ 0 \end{bmatrix} \quad .$$

The matrix

$$\mathcal{P}_{hor} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

is called the Jones matrix but we will prefer to call it as a *projection operator*.

For the case of the reflected wave we have

$$\begin{bmatrix} A \\ B \end{bmatrix} \longrightarrow \begin{bmatrix} 0 \\ B \end{bmatrix} \qquad \text{reflected wave.}$$

So the reflected wave is also a projection of the initial wave. The projection operator for the reflected wave is then

$$\mathcal{P}_{ver} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad .$$

1.4.1 Arbitrary linear polarization

What if the polarizer is rotated by an angle θ and the same wave is incident on it? Let $\hat{\mathbf{n}} = \cos\theta\hat{\mathbf{x}} + \sin\theta\hat{\mathbf{y}}$. The transmitted wave will be polarized along $\hat{\mathbf{n}}$. Let us define the unit vector in the perpendicular direction by $\hat{\mathbf{m}} = -\sin\theta\hat{\mathbf{x}} + \cos\theta\hat{\mathbf{y}}$. To determine the Jones vector of the transmitted and reflected waves, we need to express the electric field of the incident wave in the basis $\hat{\mathbf{n}}$ and $\hat{\mathbf{m}}$. In other words, we need to find numbers A' and B' in the equation

$$A\hat{\mathbf{x}} + B\hat{\mathbf{y}} = A'\hat{\mathbf{n}} + B'\hat{\mathbf{m}} \quad .$$

There are different ways of solving this equation. For us, the simplest method is the best. Just remember that $\hat{\mathbf{n}}$ and $\hat{\mathbf{m}}$ are perpendicular to each other: Their dot products vanish.

$$\begin{aligned} A' &= \hat{\mathbf{n}} \cdot (A\hat{\mathbf{x}} + B\hat{\mathbf{y}}) = A\cos\theta + B\sin\theta \\ B' &= \hat{\mathbf{m}} \cdot (A\hat{\mathbf{x}} + B\hat{\mathbf{y}}) = -A\sin\theta + B\cos\theta \end{aligned}$$

However, we would like to do the same algebra with matrices. Let us define u_n and u_m as

$$u_n = \begin{bmatrix} n_x \\ n_y \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \quad , \quad u_m = \begin{bmatrix} m_x \\ m_y \end{bmatrix} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} \quad .$$

The matrix u_n is the one that appears in the Jones matrix of a wave linearly polarized along $\hat{\mathbf{n}}$ and u_m is the one that appears for linear polarization along $\hat{\mathbf{m}}$. What we want to do is to consider the wave as a superposition of two waves with linear polarization along $\hat{\mathbf{n}}$ and $\hat{\mathbf{m}}$. In other words

$$W = \begin{bmatrix} A \\ B \end{bmatrix} = A' u_n + B' u_m = A' \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} + B' \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} \quad .$$

The method of finding A' and B' is the same. The column matrices u_n and u_m satisfy the orthonormality relations.

$$u_n^\dagger u_n = u_m^\dagger u_m = 1, \quad u_n^\dagger u_m = u_m^\dagger u_n = 0 \quad .$$

Since, this kind of products will occur frequently in quantum mechanics, we will have a special notation for them. For any two column vectors (say u and v) the *inner product* or the *scalar product* is defined as

$$\langle u|v\rangle = u^\dagger v \quad .$$

If we have

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad , \quad v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad \text{then} \quad \langle u|v\rangle = u^\dagger v = \begin{bmatrix} u_1^* & u_2^* \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = u_1^* v_1 + u_2^* v_2 \quad .$$

Later on we will learn certain mathematical properties of such products but at this point we don't need to know much. For A' and B' we have

$$\begin{aligned} A' &= \langle u_n|W\rangle = \begin{bmatrix} \cos\theta & \sin\theta \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = A\cos\theta + B\sin\theta \\ B' &= \langle u_m|W\rangle = \begin{bmatrix} -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = -A\sin\theta + B\cos\theta \end{aligned}$$

Therefore, the transmitted and reflected waves are

$$\begin{aligned} \begin{bmatrix} A \\ B \end{bmatrix} &\longrightarrow (A\cos\theta + B\sin\theta) \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \qquad \text{transmitted wave.} \\ \begin{bmatrix} A \\ B \end{bmatrix} &\longrightarrow (-A\sin\theta + B\cos\theta) \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} \qquad \text{reflected wave.} \end{aligned}$$

We succeeded in turning a simple physical experiment into a complicated mathematical problem! But please note that what we have done with matrices is exactly the same thing that we have done with vectors in the previous page. If there is anything new in here, it is the possibility of using complex numbers in the components of vectors.

There is one advantage of using the notation in here. The projection operator for finding the transmitted wave's representation can be obtained in a simple way. If the Jones vector of the incident wave is denoted by W, the projection is

$$W \longrightarrow \langle u_n|W\rangle u_n = (u_n^\dagger W) u_n = u_n (u_n^\dagger W) = (u_n u_n^\dagger) W$$

where $u_n u_n^\dagger$ is a 2×2 matrix, and in the expression above we have that matrix multiplied with the incident wave's Jones vector. Therefore, the corresponding Jones operator for transmitted wave is

$$\mathcal{P}_{\hat{\mathbf{n}}} = u_n u_n^\dagger = \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix} \quad ,$$

and for the reflected wave is

$$\mathcal{P}_{\hat{\mathbf{m}}} = u_m u_m^\dagger = \begin{bmatrix} \sin^2\theta & -\cos\theta\sin\theta \\ -\cos\theta\sin\theta & \cos^2\theta \end{bmatrix} \quad .$$

Please note that the sum of the two projection operators is equal to the identity matrix

$$\mathcal{P}_{\hat{\mathbf{n}}} + \mathcal{P}_{\hat{\mathbf{m}}} = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

which expresses the fact that the wave is either transmitted or reflected, and if you superpose both of the waves again you will obtain the original wave. Also, note that the direction of $\hat{\mathbf{n}}$ is not important. A polarizer with axis along $-\hat{\mathbf{n}}$ ($\theta \to \theta + \pi$) would produce the same results.

1.5 Energy

The total energy carried by the wave is given by

$$U = \frac{1}{8\pi}\int(\vec{\mathbf{E}}^2 + \vec{\mathbf{B}}^2)d^3\vec{\mathbf{r}} = \frac{1}{4\pi}\int\vec{\mathbf{E}}^2 d^3\vec{\mathbf{r}} = \frac{1}{8\pi}(E^2_{x,max} + E^2_{y,max})\mathcal{V}$$

where $E_{x,max}$ and $E_{y,max}$ are the maximum value of the electric fields along x and y respectively, and $\mathcal{V}$ is the average volume occupied by the wave. If

$$W = \begin{bmatrix} A \\ B \end{bmatrix}$$

is the Jones vector for a wave then the total energy is

$$U = \frac{1}{8\pi}(|A|^2 + |B|^2)\mathcal{V} = \frac{1}{8\pi}W^\dagger W\mathcal{V} \quad .$$

We are going to draw two important points from this expression. The first is that, upon meeting a polarizer, the total energy is divided into two. For the case of a polarizer with polarization axis along x, the energy carried by the transmitted wave is

$$U_{trans} = \frac{1}{8\pi}|A|^2\mathcal{V}$$

and the energy carried by the reflected (or absorbed) wave is

$$U_{refl} = \frac{1}{8\pi}|B|^2\mathcal{V} \quad .$$

It can be seen that the total energy is conserved ($U = U_{trans} + U_{refl}$). We can summarize this by saying that the fraction

$$\frac{|A|^2}{|A|^2 + |B|^2}$$

of the energy is carried by the transmitted wave and the rest by the reflected.

The second point is that the magnitude of the electric field amplitude depends on the volume occupied by the wave. This amplitude can take on any values; it can be large or small depending on the volume. This is even if there is only one photon. Therefore the magnitude of the fields is unrelated to the polarization state of the wave. This gives us sufficient reason to eliminate references to the electric field in the matrix representation of polarization.

Problem 1. A linearly polarized light is incident on a linear polarizer and their polarization directions makes angles θ_L and θ_P respectively. Show that the fraction of energy transmitted is $\cos^2(\theta_L - \theta_P)$.

Problem 2. Suppose that a right circularly polarized wave is incident on a linear polarizer. Assume that the polarization axis of the polarizer makes an angle θ with the x axis. Which fraction of energy is transmitted? Which fraction is absorbed? What happens with left circularly polarized light?

1.5.1 Orthogonality

There are also circular polarizers. A right circular polarizer is transparent to right-handed waves and absorbs/reflects left-handed waves. It is also possible to construct more general elliptical polarizers. Can we say anything about them?

Suppose that a particular polarizer is transparent to waves with polarization u_1 and reflects all waves with polarization u_2. Using the conservation of energy argument, we can show that u_1 and u_2 are orthogonal to each other. In other words, u_1 and u_2 satisfy the mathematical relation

$$\langle u_1|u_2\rangle = u_1^\dagger u_2 = 0 \quad .$$

To show this first consider how the polarizer acts for certain polarizations. If a wave with Jones vector Au_1 is incident on the polarizer, it will be completely transmitted. On the other hand if a wave with

Jones vector Bu_2 is incident, then it will be completely reflected. Let us consider what happens if the Jones vector, W, of the incident wave is arbitrary. We can think of the incident wave as a superposition of two waves with polarizations u_1 and u_2. Hence, $W = Au_1 + Bu_2$ for some numbers A and B. Then the transmitted and reflected waves will have the Jones vectors Au_1 and Bu_2 respectively.

On the other hand, the conservation of energy requires

$$\begin{aligned} U &= U_{trans} + U_{refl} \\ \frac{1}{8\pi} W^\dagger W \mathcal{V} &= \frac{1}{8\pi}(Au_1)^\dagger (Au_1)\mathcal{V} + \frac{1}{8\pi}(Bu_2)^\dagger (Bu_2)\mathcal{V} \\ (Au_1 + Bu_2)^\dagger (Au_1 + Bu_2) &= |A|^2 u_1^\dagger u_1 + |B|^2 u_2^\dagger u_2 \\ A^* B u_1^\dagger u_2 + AB^* u_2^\dagger u_1 &= 0 \end{aligned}$$

Since this expression has to be true for all possible values of A and B we give special values to them and try to reduce the equation further. If A and B are both real we have $u_1^\dagger u_2 + u_2^\dagger u_1 = 0$. On the other hand if A is real and B is purely imaginary we have $u_1^\dagger u_2 - u_2^\dagger u_1 = 0$. The orthogonality relation

$$u_1^\dagger u_2 = 0 \tag{8}$$

then follows.

This is a very strong relation. It says that if a polarizer transmits horizontal waves, it should reflect vertical ones. If it transmits right circular waves, it should reflect left circular ones.

Problem 3. Verify that

$$u_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix} \quad \text{and} \quad u_2 = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

are orthogonal to each other.

Problem 4. A polarizer transmits completely any wave with Jones vector

$$W = A \begin{bmatrix} 2 \\ 1+i \end{bmatrix} \quad .$$

(This is usually called as an elliptical polarization.) If light with some other polarization is completely reflected, then what would be the Jones vector for this case?

1.5.2 What is relevant for the polarization state?

Let us say a few words about which information is relevant for the polarization state in the Jones vector notation. First of all we have said above that the magnitude of the electric field depends only on the total energy carried by the wave and the spatial extent of the wave. But these have nothing to do with polarization. A horizontally polarized wave might carry two times more energy (and in this case the electric fields will be $\sqrt{2}$ times larger) but it is still horizontally polarized. Same for waves occupying different volumes. However, in the Jones vector notation

$$W = \begin{bmatrix} A \\ B \end{bmatrix} \quad ,$$

$|A|$ and $|B|$ gives the maximum values of the x and y components of the electric field vector respectively. We can drop that information about the magnitude of the electric fields from our notation. This means that, we can multiply both components of W by the same real number and still carry the same information about polarization. In other words, both of

$$W = \begin{bmatrix} A \\ B \end{bmatrix} \quad \text{and} \quad W' = \begin{bmatrix} cA \\ cB \end{bmatrix} \quad ,$$

where c is a real number, represent the same polarization state. The only important point is that both components should be multiplied with the same number, since the relative magnitudes of E_x and E_y is important for the polarization.

A second point is about the complex phase of the components. Let θ_1 and θ_2 represent the phases of the two components of the Jones vector.

$$W = \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} |A|e^{i\theta_1} \\ |B|e^{i\theta_2} \end{bmatrix} \quad .$$

This means that the electric fields are

$$\begin{aligned} E_x &= |A|\cos(kz - \omega t + \theta_1) \quad , \\ E_y &= |B|\cos(kz - \omega t + \theta_2) \quad . \end{aligned}$$

It can be seen that the exact value of the phases depends on where you have chosen your origin (the point with $z = 0$) and when you have started your clock (the moment when $t = 0$). Since these choices are up to you, there is some irrelevant information in here as well. For example, if you have started your clock at time t_1, the phases you have to use for your new choices will be

$$\begin{aligned} \theta_{1,new} &= \theta_1 - \omega t_1 \quad , \\ \theta_{2,new} &= \theta_2 - \omega t_1 \quad . \end{aligned}$$

But the wave is still the same wave. This implies that you can shift both phases by the same amount and still describe the same wave! This is equivalent to multiplying both components of the Jones vector by the same phase factor ($\exp(-i\omega t_1)$ in here). Same as above, it is important to change both phases by the *same* amount. The relative phase difference of the two components (i.e., $\theta_2 - \theta_1$) is important.

Together with the above, this implies that you can multiply both components of a Jones vector without changing the information about the polarization state. As a result, all the following vectors

$$\begin{bmatrix} A \\ 0 \end{bmatrix} \quad , \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad , \quad \begin{bmatrix} -1 \\ 0 \end{bmatrix} \quad , \quad \begin{bmatrix} i \\ 0 \end{bmatrix}$$

describe a horizontally polarized wave while all of the following

$$\begin{bmatrix} A \\ -iA \end{bmatrix} \quad , \quad \begin{bmatrix} 1 \\ -i \end{bmatrix} \quad , \quad \begin{bmatrix} i \\ 1 \end{bmatrix} \quad , \quad \begin{bmatrix} 1+i \\ 1-i \end{bmatrix}$$

describe a right circularly polarized wave.

These simplifications are carried out whenever possible. However, when you are dealing with the superposition of two different waves, the relative magnitudes of the electric field vectors and the relative phases of the two waves become important. So, for such cases, either you don't carry out this simplification, or you do so for both waves at the same time. Same mathematical simplifications will be possible for the quantum mechanical description of the polarization state of a photon, but for a different reason.

2 Polarization State of a Photon

Suppose that we are experimenting with a polarized electromagnetic wave, but we decreased the intensity of the beam to extremely small values so that at any given time there can be at most one photon within the experimental set-up. The question is "how does the photon behave when it meets a polarizer?" The answer is quantum-mechanical obviously, and we have to make sure that the answers that we give should be consistent with the behavior of classical waves. That is, if we repeat the same experiment over and over again for millions of photons and sum up the results, the final result should be the same as the classical theory predicts. This is the Correspondence Principle.

Consider first a photon incident on a linear polarizer with polarization axis oriented along x. A classical wave incident on this polarizer would be split into two with appropriate polarizations. What would happen to the photon? The one thing that we know is that the photon is indivisible. If the frequency is $f = \omega/2\pi$, the photon has the energy $hf = \hbar\omega$ and there is no way that it can carry a smaller energy. Therefore, the photon either passes the polarizer (it is transmitted) or it is reflected. It cannot be divided into two.

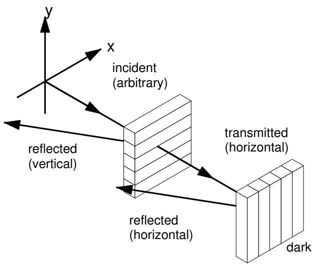


Figure 1: The light that is transmitted by the first polarizer is definitely reflected by the second polarizer. Therefore, the photons should do the same.

(To be strict we are forcing the Copenhagen interpretation in here. Think of it like this. Suppose that we detect the photon and find out where it has gone. In other words we measure its position. Then the answer would be either 'reflected' or 'transmitted', but not both.)

Since many photons would make up a classical wave, then we can say that some photons are transmitted and some are reflected. That is if the classical wave has an incidence polarization such that it is divided into a transmitted and reflected beams, then some of the photons have to be transmitted and some have to be reflected. Therefore, when the photon starts to pass the polarizer, some mechanism starts to take place and right at that moment it is decided in which way the photon should go. That mechanism has to be un-deterministic. We cannot predict before the incidence in which way the photon goes.

Next, let us find the probability of transmission and reflection. Suppose that the classical wave has the Jones vector

$$W = \begin{bmatrix} A \\ B \end{bmatrix} \quad .$$

The classical theory tells us that the fraction

$$\frac{|A|^2}{|A|^2 + |B|^2}$$

of the energy has to be transmitted when the polarizer is oriented along x. Since all photons have the same energy, this should be equal to the probability that the photons are transmitted

$$P_{trans} = \frac{|A|^2}{|A|^2 + |B|^2} \quad .$$

The reflection probability, then, has to be

$$P_{refl} = 1 - P_{trans} = \frac{|B|^2}{|A|^2 + |B|^2} \quad .$$

2.0.3 The state after

Now, suppose that we have sent a photon to the polarizer and found that it is transmitted. Can we say that the photon now is horizontally polarized? We can rephrase this question as "Is a horizontally polarized wave formed only by horizontally polarized photons or can there be a few photons with other polarizations?" To answer that question we can appeal to the Correspondence Principle again. Place another polarizer in front of the transmitted beam with polarization direction along y. For a classical wave, the wave is split up in the first polarizer. The transmitted wave then has horizontal polarization and therefore it cannot pass the second polarizer. Since no energy can pass the second polarizer, no photons can pass it as well. Then the answer is obvious: When the photon passes the first polarizer, its polarization state changes to horizontal polarization. Similarly, if the photon is reflected by the first polarizer, its polarization state changes to the vertical.

This is the Copenhagen interpretation of quantum mechanics. *When a measurement is carried out on a state, (1) only one of the few possible results are found probabilistically and (2) the state collapses to* *another state consistent with the experimental result.* We are going to make this expression a little bit more exact in a moment. But, let us first note that the process of sending a photon to a polarizer is a measurement. We are "measuring" the polarization state of the photon. However, this measurement cannot tell us everything about the initial polarization state of the photon. It tells us either that the photon has vertical polarization or horizontal polarization. Moreover, after the measurement, the photon changes its polarization to either one of these. The initial polarization state of the photon has collapsed into either horizontal or vertical polarization state. Moreover, after the measurement is done, we can not say any further things about the initial state. The collapse has cleared every information about the initial state.

2.1 The State Vector

Before going further, let us correct our notation and eliminate all references to the electric field. If the Jones vector is

$$W = \begin{bmatrix} A \\ B \end{bmatrix}$$

where A and B are complex-numbers in electric field units, we define the quantum-mechanical *state* (or *state vector*) of a photon in the wave as

$$\chi = \frac{1}{\sqrt{|A|^2 + |B|^2}} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \quad .$$

Here, a and b are dimensionless complex numbers. Also, the state χ is normalized

$$\chi^\dagger \chi = |a|^2 + |b|^2 = 1 \quad .$$

The probabilities of transmission and reflection are simply

$$P_{trans} = |a|^2 \qquad , \qquad P_{refl} = |b|^2 \quad .$$

To which state does the collapse occur? In terms of W, the collapse for the transmitted wave is

$$\begin{bmatrix} A \\ B \end{bmatrix} \longrightarrow \begin{bmatrix} A \\ 0 \end{bmatrix} \quad .$$

However, this cannot be the case for a photon. In the notation above, the final state has smaller electric field, and therefore it has smaller energy. (Passing a polarizer cannot change the volume occupied by the photon). Since the photon should still have the same energy, after passing the polarizer, the electric field is not simply projected, it is projected and renormalized to the appropriate magnitude

$$\begin{bmatrix} A \\ B \end{bmatrix} \longrightarrow \begin{bmatrix} \sqrt{|A|^2 + |B|^2} \\ 0 \end{bmatrix} \quad .$$

This is an important difference between the quantum and classical pictures. In any case, in quantum mechanics we use normalized states, so that in terms of χ the collapse is

$$\chi = \begin{bmatrix} a \\ b \end{bmatrix} \longrightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad .$$

Note that we don't say anything about the phase after the collapse. Unless you are doing an experiment where superposition is involved that phase is unimportant. For this reason, in the following we will ignore all such phase factors.

Finally we can rephrase the Copenhagen interpretation for this experiment. *If the polarization state of a photon is*

$$\chi = \begin{bmatrix} a \\ b \end{bmatrix}$$

and it is incident on a polarizer with its axis along x, then with probability $P_{trans} = |a|^2$ the photon is transmitted and its state collapses to

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad ,$$

or it is reflected with probability $P_{refl} = |b|^2$ and its state collapses to

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad .$$

After the experiment, the photon loses all information about its initial state.

2.2 Other Polarization Directions

Suppose now that the polarizer is tilted so that its polarization axis makes angle θ with x-axis. A photon in the state

$$\chi = \begin{bmatrix} a \\ b \end{bmatrix}$$

is incident on the polarizer. What is the probability of transmission? Again, we define $\hat{\mathbf{n}} = \cos\theta\hat{\mathbf{x}} + \sin\theta\hat{\mathbf{y}}$ and $\hat{\mathbf{m}} = -\sin\theta\hat{\mathbf{x}} + \cos\theta\hat{\mathbf{y}}$. All photons that pass the polarizer will have their polarizations along $\hat{\mathbf{n}}$. That means that they will collapse to the state

$$\chi_{\hat{\mathbf{n}}} = \begin{bmatrix} n_x \\ n_y \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \quad .$$

On the other hand, all photons that are reflected will be polarized along $\hat{\mathbf{m}}$ and this means that they will collapse to

$$\chi_{\hat{\mathbf{m}}} = \begin{bmatrix} m_x \\ m_y \end{bmatrix} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} \quad .$$

We should then, think of the initial state of the photon, χ, to be a linear superposition of these two states.

$$\chi = a'\chi_{\hat{\mathbf{n}}} + b'\chi_{\hat{\mathbf{m}}} \quad . \tag{9}$$

Finding the numbers a' and b' is a similar simple problem due to orthonormality of $\chi_{\hat{\mathbf{n}}}$ and $\chi_{\hat{\mathbf{m}}}$.

$$\begin{aligned} a' &= \chi_{\hat{\mathbf{n}}}^\dagger \chi = \langle \chi_{\hat{\mathbf{n}}} | \chi \rangle \\ b' &= \chi_{\hat{\mathbf{m}}}^\dagger \chi = \langle \chi_{\hat{\mathbf{m}}} | \chi \rangle \end{aligned}$$

The question was what are the probabilities of transmission and reflection. To answer that, we have to think of the Eqn. (9) as a superposition of two classical waves forming another classical wave. Since, out of a total energy, only the fraction

$$\frac{|a'|^2}{|a'|^2 + |b'|^2}$$

is transmitted, the transmission probability of the photon has to be

$$P_{trans} = |a'|^2$$

where we have used the fact that $|a'|^2 + |b'|^2 = 1$. Similarly, the reflection probability is

$$P_{refl} = 1 - P_{trans} = |b'|^2 \quad .$$

For this reason, the complex numbers a' and b' are called *probability amplitudes.*

Problem 5. A photon linearly polarized with an angle θ with x axis is incident on a polarizer with its axis oriented along a direction with angle θ'. What is the probability that the photon is transmitted? Does the answer change if we make the changes $\theta \to \theta + \pi$ or $\theta' \to \theta' + \pi$? What is the answer for $\theta = \pi/4$ and $\theta' = 0$?

Problem 6. A photon is in a right circularly polarized state and is incident on a linear polarizer with its axis oriented along a direction with angle θ'. What is the probability of transmission? Does the answer depend on θ'?

2.3 General Case

Consider a polarizer which transmits waves with polarizations u_1 and blocks the ones with u_2. We assume that u_1 and u_2 are normalized

$$u_1^\dagger u_1 = \langle u_1 | u_1 \rangle = 1 \qquad , \qquad u_2^\dagger u_2 = \langle u_2 | u_2 \rangle = 1 \quad .$$

We also know that conservation of energy requires u_1 and u_2 to be orthogonal

$$u_1^\dagger u_2 = \langle u_1 | u_2 \rangle = 0 \quad .$$

As a result we have an orthonormal set

$$\langle u_\ell | u_j \rangle = \delta_{\ell,j} \qquad (\ell, j = 1, 2)$$

If a photon is in a normalized state χ, then

$$\chi = a_1 u_1 + a_2 u_2 \quad ,$$

where

$$a_j = \langle u_j | \chi \rangle \quad .$$

The probability of photon being transmitted is

$$P_{trans} = |a_1|^2 = |\langle u_1 | \chi \rangle|^2$$

and the state χ collapses to u_1. The reflection probability is

$$P_{refl} = |a_2|^2 = |\langle u_2 | \chi \rangle|^2$$

and the state collapses to u_2.
Problem 7. For the case of a right circular polarizer, u_1 and u_2 are

$$u_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} \qquad , \qquad u_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} \quad .$$

Show that for any linearly polarized photon, probability of transmission from a right circular polarizer is independent of the polarization angle.