Lecture 1: Light Polarization, Wednesday, Aug. 31

Quantum mechanics is perhaps one of the most brilliant discoveries the human kind ever made. The mathematical tools that it uses are in fact familiar in classical wave mechanics, which is a good reason for us to start the course by discussing the polarization of light. What is non-trivial, however, is that the quantity that one directly deals with, the so-called \textit{probability amplitude}, is NOT a physical observable. Therefore, it is really a huge stretch of imagination to use the known machinery of wave mechanics to something not directly observable or seemingly pure-mathematical (the so-called “wave” behavior of the microscopic systems), and yet the whole thing remains sensible and passes every experimental tests that physicists have come up with so far!

Consider a plane light wave travelling in the \( z \)-direction, with wave vector \( k \) and frequency \( \omega \). Its electric field may be written as

\[
\vec{E} = \vec{E}_0 e^{-i(\omega t - k z)} .
\]  

(1)

Clearly \( \vec{E}_0 \) must be in the \( x - y \) plane (light-wave is transverse wave!), and \( \vec{E}_0 \) can be complex if one adopts the convention that only the real part of \( \vec{E} \) is physical. We will ignore the magnetic field which is not essential for our purpose here.

If \( \vec{E}_0 \) is along the \( x \)-direction: \( \vec{E}_0 = E_0 x \vec{e}_x \), we say the light is \textit{linearly} polarized along the \( x \)-direction. On the other hand, if \( \vec{E}_0 = E_0 y \vec{e}_y \), we say the light is \textit{linearly} polarized along the \( y \)-direction. \([\vec{e}_x^2 = \vec{e}_y^2 = 1, \vec{e}_x \cdot \vec{e}_y = 0]\)

From experience, these two polarizations are independent of each other, and are complete in the sense that any other polarization can be expressed as a vector sum of the two,

\[
\vec{E}_0 = E_{0x} \vec{e}_x + E_{0y} \vec{e}_y .
\]  

(2)

Therefore we may postulate that all polarizations form a two-dimensional \textit{complex linear vector space}.  

1
Let us argue that the complexity of $E_0x$ and $E_0y$ is really necessary. Suppose both of them are real, the resulting electric field makes a $\theta$ angle relative to the $x$ axis, with $\theta = \tan^{-1}(E_{0y}/E_{0x})$. This again is a linearly polarized light. One can also add a constant phase $e^{i\phi}$ to both components, the conclusion remains the same: Light polarization does NOT depend on the overall phase of the two components.

On the other hand, light polarization is characteristically different if one of the components has a relative phase to the other. In this case, one obtains what is called the elliptic polarization, in which the electric field vector $\text{Re} \bar{E}$ rotates in the $x$-$y$ plane with its tip tracing out an ellipse (see J. D. Jackson, *Classical Electrodynamics*, 2nd edition, p. 276, Fig. 7.4). [The center of the ellipse is at the origin of the coordinates.] If $|E_{0x}| = |E_{0y}|$ and the phase difference is $\pi/2$, the elliptically polarized light becomes circularly polarized. In particular,

$$\bar{E}_0 = E_0(\vec{e}_x + i\vec{e}_y)$$

is a right-handed polarized light in the sense that the direction of rotation (counterclockwise when facing the beam) and the direction of the light propagation obey the right-handed rule (in the literature of optics, it is actually called left-circularly polarized). If $i \rightarrow -i$, we get a left-handed (right-circularly, or clockwise) polarized light.

To see further the use of the complex vector space, let us calculate the intensity of the light $I$. It is easy to show that

$$I \sim |E_{0x}|^2 + |E_{0y}|^2 = E_{0x}^* E_{0x} + E_{0y}^* E_{0y}$$

The right-hand side can be considered as the norm of the complex vector $\bar{E}_0$, which is a natural concept in a complex vector space.

To distinguish the 2D complex vector space of the light polarization from the ordinary 2D space, let us consider the math structure of the former. Choose the $x$ and $y$ linearly-polarized lights as our basis, we introduce the Dirac’s ket notation

$$|e_x\rangle, \ |e_y\rangle$$

[Although Dirac’s notation was invented to discuss quantum mechanics, there is no reason that one cannot use it here for the light polarization. In fact, it can be used for vectors in any vector space.] Then, Eq. (2) can be written as (omitting 0 and vector sign)

$$|E\rangle = E_x|e_x\rangle + E_y|e_y\rangle$$

2
We also introduce the matrix notation,

\[ |E\rangle \sim \begin{pmatrix} E_x \\ E_y \end{pmatrix} \] (7)

Although it is tempting to think of \( |e_x\rangle \) and \( |e_y\rangle \) as the unit vectors in the x and y directions in the x-y plane, the complex vector space can also have a new type of bases, such as

\[ |L\rangle = \frac{1}{\sqrt{2}} (|e_x\rangle - i|e_y\rangle) \]

\[ |R\rangle = (-1)^{1/2} (|e_x\rangle + i|e_y\rangle) \] (8)

(Where - sign is optional) which represent left and right-handed polarizations and are clearly not vectors in the x-y plane. Certainly any polarization can be written as a linear superposition of the left and right-handed polarizations

\[ |E\rangle = E_-|L\rangle + E_+|R\rangle . \] (9)

To define the concept of norm and to calculate linear superposition easily, we introduce the conjugate bra vector for every ket vector,

\[ \langle e_x|, \langle e_y|, \langle L|, \langle E|,... \] (10)

In particular

\[ \langle E| = \langle e_y|E_x^* + \langle e_y|E_y^* \] (11)

which admits a matrix form,

\[ \langle E| \sim (E_x^*, E_y^*) \] (12)

Then it is natural to introduce the inner product between a pair of bra and ket vectors

\[ \langle E|E'\rangle = c, \quad \langle E'|E\rangle = c^* \] (13)

where \( c \) is an complex number. The basis vectors are orthonormal in the sense

\[ \langle e_i|e_j\rangle = \delta_{ij} \] (14)

which is a mathematical statement that two basis polarizations are physically independent.
Now it is easy to see that the coefficients of the linear superposition are just the inner product between the polarization ket and the basis bras,

\[ E_i = \langle e_i | E \rangle \]  \hspace{1cm} (15)

Moreover, if we don’t care about the intensity of the light, we can assume the polarization ket is normalized

\[ \langle E | E \rangle = |E_x|^2 + |E_y|^2 = 1 \]  \hspace{1cm} (16)

Since the overall phase of the ket is irrelevant, the light polarization is determined by two real parameters (relative strength and phase between the x and y components).

The linear superposition principle is a hallmark of wave physics, and is physically quite striking. It says that any polarization can be viewed as a coherent superposition of two basis polarizations, here “coherent” means that the relative phase is fixed. Therefore, one can, for example, decompose physically a light into two components with orthogonal polarizations. Likewise, one can add together two or more polarizations with fixed phases to obtain a single polarization. These principles of course are similar to decomposing and adding positions and velocity vectors in classical mechanics. We will consider some of the implications soon.

Having established the state space for the light polarizations, we are ready to consider measurements. A polarizer is an apparatus allowing a particular polarization component going through, but blocking the orthogonal component. For example, an X polarizer allows the \( E_x \vec{e}_x \) component of the electric field going unchanged, but eliminating \( E_y \vec{e}_y \) component. In the 2D complex vector space, we use an operator \( \hat{P}_x \) to represent the polarizer, and we have

\[ \hat{P}_x |E\rangle = E_x |e_x\rangle \]  \hspace{1cm} (17)

In fact, we can use a 2 matrix to represent the polarizer

\[ \hat{P}_x = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \]  \hspace{1cm} (18)

which is a hermitian matrix in the sense that

\[ A^\dagger = A, \quad A_{ij} = A_{ji}^* \]  \hspace{1cm} (19)
The Y polarizer can also be represented by a hermitian matrix,

\[ \hat{P}_y = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \]  

(20)

And it is easy to see that \( \hat{P}_x + \hat{P}_y = 1 \), a unit matrix. In fact, all apparatus can be represented by a hermitian matrix in the state space.

Hermitian matrices have real eigenvalues, and their eigenvectors corresponding to distinct eigenvalues are orthogonal to each other. For example, \( \hat{P}_x \) has two eigenvalues: 1 and 0. The orthogonal eigenvectors are \( |e_x\rangle \) and \( |e_y\rangle \), respectively. After the light passing through the X polarizer, the component corresponding to eigenvalue 1 is allowed to go through, that corresponding to 0 is blocked. Therefore, after measurement the light reduces to the eigenstate \( |e_x\rangle \) of the polarizer.

The hermitian operator corresponding to the X polarizer can also be written as

\[ \hat{P}_x = |e_x\rangle\langle e_x| \]  

(21)

As we shall see, under hermitian conjugation, \( |\psi\rangle\langle \eta| \) becomes \( |\eta\rangle\langle \psi| \).

It can be shown that if the light is polarized in the x direction, and it passes through a polarizer oriented in \( \theta \) angle relative to the x axis, the light intensity is reduced by a factor of \( \cos^2 \theta \). This is called Malus’s law in optics.

Finally, let us discuss a phenomenon which seems peculiar. Consider a light polarized in the x direction. After passing through a Y polarizer the light intensity becomes 0, obviously. However, if one applies a 45° polarizer before letting it goes through the Y polarizer, then there is light coming out of the latter. The total light going through the combined 45° and Y polarizer is 1/4 of the original. The explanation for this phenomenon is the superposition principle.