Polarization

Wesley C. Campbell

July 2, 2019

1 Linear Polarization

For Physics 180Q, we will often be approximating the light that comes from a laser as being:

1. composed of infinite plane-waves
2. perfectly polarized
3. perfectly monochromatic.

These three approximations break down in various interesting ways, all of which we will explore in this class. However, for the present moment, we will use these simplifications to write the electric field of a linearly-polarized laser beam propagating in the $+z$-direction as

$$\mathbf{E}_{\mathrm{x}}(z,t) = \hat{\mathbf{x}}\,\mathrm{I\!R}\Big\{\mathcal{E}_0\,\mathrm{e}^{i(kz-\omega t)}\Big\} \tag{1}$$

if it is linearly polarized along x and

$$\mathbf{E}_{\mathrm{y}}(z,t) = \hat{\mathbf{y}}\,\mathrm{I\!R}\Big\{\mathcal{E}_0\,\mathrm{e}^{i(kz-\omega t)}\Big\} \tag{2}$$

if it is linearly polarized along y. Before getting into the details, let's establish some conventions.

The "$\mathrm{I\!R}\{\cdots\}$" above is my notation for taking the real part of an expression, but I am not going to bother to write that very frequently. You should assume that from now on, we're always going to be taking the real part of an expression to get the actual electric field. Note that this is the same thing as dividing by 2 after adding the complex conjugate of these expressions, so be aware of that when you get into the quantum mechanics of photons.

A similar statement will apply, incidentally, to the $\mathrm{e}^{i(kz-\omega t)}$ term – this thing will be riding along with lots of expressions in a pretty boring way, so please don't get too worried if I stop writing it in situations where it doesn't play an important role.

Now, you know that there are other types of polarization than just those described by Eq. (1) and (2). Linearly-polarized light that propagates along $+z$ can be polarized in *any* direction in the xy plane, and this is not to mention the more-exotic polarization states known as "circular" and "elliptical." We will cover all of this, so not to worry – we're going to start with the basics and build in a fully-general framework on which to hang that stuff later. This is another way for me to say that the two forms of polarization above form a *complete basis* for arbitrary polarization states of the light, a fact we will utilize shortly.

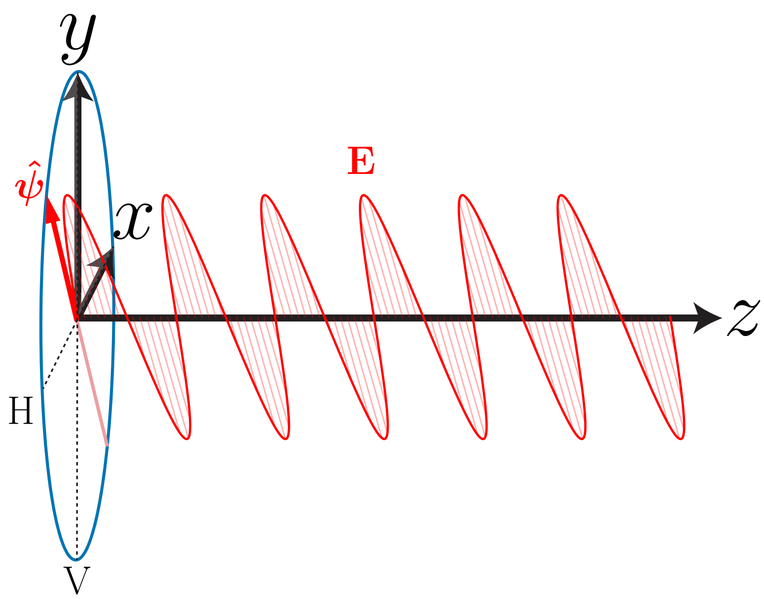

Figure 1: Coordinate system for defining the polarization of light propagating in the $+z$ direction. We will take the xz-plane to be horizontal (H) and the yz-plane to be vertical (V).

1.1 Jones vectors

Okay, it seems clear that for a laser beam like this, the polarization itself can be described by the unit vector out in front. Let's call this thing $\hat{\boldsymbol{\psi}}$ in anticipation of the fact that it's going to look a lot like a quantum-mechanical state vector $|\psi\rangle$. For the moment, the important thing to know is just that $\hat{\boldsymbol{\psi}}$ is a unit vector that points in some direction in the xy plane, and that we will eventually allow it to become complex.

Using this, we can write

$$\mathbf{E} = \hat{\boldsymbol{\psi}}\,\mathcal{E}_0\,\mathrm{e}^{i(kz-\omega t)}. \tag{3}$$

Now, notice that if I compare light with $\hat{\boldsymbol{\psi}} = \hat{\mathbf{x}}$ to light with $\hat{\boldsymbol{\psi}} = -\hat{\mathbf{x}}$, the only difference can be interpreted as a phase shift of $-1 = \mathrm{e}^{i\pi}$. In contrast, $\hat{\boldsymbol{\psi}} = \hat{\mathbf{x}}$ and $\hat{\boldsymbol{\psi}} = \hat{\mathbf{y}}$ are clearly different beasts. So since we have stipulated that (for the moment) the laser beam is propagating along $+z$, we can orient our laser beam and its associated coordinate system such that anything polarized along either x or $-x$ has *horizontal* polarization (denoted by H), and anything polarized along $\pm y$ has *vertical* polarization (denoted by V), as shown in Fig. 1.

Since this vector $\hat{\boldsymbol{\psi}}$ lives in the xy plane, it has two components, and we will always be able to write it as a column vector,

$$\hat{\boldsymbol{\psi}} = \begin{pmatrix} a_{\mathrm{H}} \\ a_{\mathrm{V}} \end{pmatrix} \tag{4}$$

where the upper component is proportional to the amplitude that is horizontally-polarized and the lower is proportional to the amplitude that is vertically polarized, subject to the normalization condition $\hat{\boldsymbol{\psi}}^{\dagger}\hat{\boldsymbol{\psi}} = |a_{\mathrm{H}}|^2 + |a_{\mathrm{V}}|^2 = 1$. So we can see that

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv \hat{\boldsymbol{\psi}}_{\mathrm{H}} \quad \text{would be pure horizontal polarization, and}$$
$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv \hat{\boldsymbol{\psi}}_{\mathrm{V}} \quad \text{would be pure vertical polarization.}$$

The column vector $\hat{\boldsymbol{\psi}}$ is called a *Jones vector*, and it turns out that it follows the exact same algebra as the quantum-mechanical spin states of a spin-1/2 particle (astute readers may have already recognized that the special basis vectors $\hat{\boldsymbol{\psi}}_{\mathrm{H}}$ and $\hat{\boldsymbol{\psi}}_{\mathrm{V}}$ above are the +1 and −1 eigenvectors of the Pauli σ_X operator).

1.2 Jones Matrices

Why is this helpful? Well, the effect of various optical elements on the polarization of the light can be represented by 2×2 matrices that transform the Jones vectors according to

$$\hat{\boldsymbol{\psi}}' = \mathbf{T}\hat{\boldsymbol{\psi}} = \begin{pmatrix} T_{\mathrm{HH}} & T_{\mathrm{HV}} \\ T_{\mathrm{VH}} & T_{\mathrm{VV}} \end{pmatrix} \begin{pmatrix} a_{\mathrm{H}} \\ a_{\mathrm{V}} \end{pmatrix}, \tag{5}$$

as shown in Fig. 2(a).

These *Jones matrices* **T** are determined by the properties of the optical elements that the light passes through, and do not depend on properties the light. This formalism is therefore the same as that of other *linear systems*, and we can apply lots of useful results right away.

For instance, if 58 polarization-transforming elements are cascaded such as shown in Fig. 2(b), the whole cascade can be represented by a single 2×2 Jones matrix,

$$\hat{\boldsymbol{\psi}}' = \mathbf{T}_{58} \cdots \mathbf{T}_2\mathbf{T}_1\hat{\boldsymbol{\psi}} = \mathbf{T}_{\mathrm{c}}\hat{\boldsymbol{\psi}}. \tag{6}$$

This is a dramatic simplification compared to dealing with 58 equations, and the resulting matrix is the same for *any* input polarization $\hat{\boldsymbol{\psi}}$.

Example 1.1 *A* polarizer *is an optical element that will transmit only a specific polarization, ideally with 100% fidelity. For instance, a horizontal polarizer will remove all of the vertically-polarized light, either by absorbing it or sending it elsewhere, as shown in Fig. 3(a). We can figure out what the Jones matrix for an ideal* H*-polarizer must be by noting its effect on a general Jones vector:*

$$\begin{pmatrix} a_{\mathrm{H}} \\ 0 \end{pmatrix} = \begin{pmatrix} T_{\mathrm{HH}} & T_{\mathrm{HV}} \\ T_{\mathrm{VH}} & T_{\mathrm{VV}} \end{pmatrix} \begin{pmatrix} a_{\mathrm{H}} \\ a_{\mathrm{V}} \end{pmatrix} \tag{7}$$

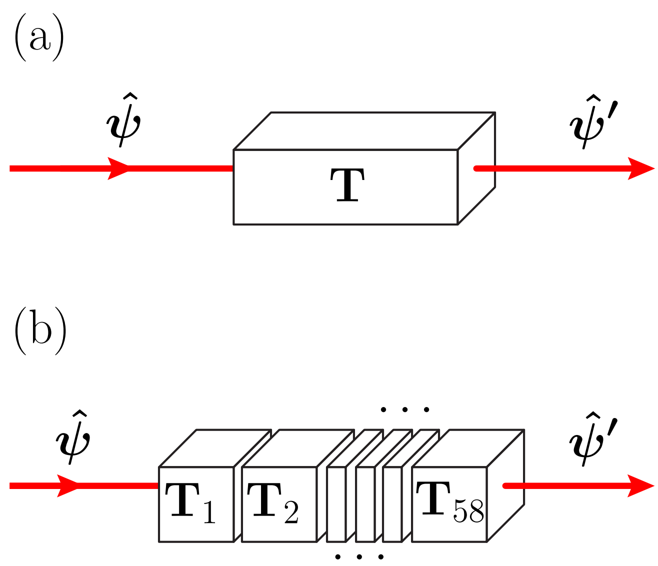

Figure 2: Linear-optics elements that transform the polarization of light can be modeled by their associated Jones matrices $\mathbf{T}$. Systems with single elements (a) and cascaded elements (b) are both represented by a single, 2×2 Jones matrix.

We can write this as a system of two equations,

$$\begin{aligned} a_{\mathrm{H}} &= T_{\mathrm{HH}}\, a_{\mathrm{H}} + T_{\mathrm{HV}}\, a_{\mathrm{V}} \\ 0 &= T_{\mathrm{VH}}\, a_{\mathrm{H}} + T_{\mathrm{VV}}\, a_{\mathrm{V}}. \end{aligned}$$

Since the elements of $\mathbf{T}$ *cannot depend upon* a_{H} *and* a_{V}*, the only way for these equations to be satisfied is if* $T_{\mathrm{HV}} = T_{\mathrm{VH}} = T_{\mathrm{VV}} = 0$ *and* $T_{\mathrm{HH}} = 1$*, which gives us the Jones matrix for a horizontal polarizer (in the* H, V *basis):*

$$\mathbf{T}_{\mathrm{Hpol.}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{8}$$

Note that this is not a unitary matrix ($|\det\{\mathbf{T}_{\mathrm{Hpol.}}\}| \neq 1$)*. Classically, that means we can't expect the action of this device to be reversible; if we send* H*-polarized light back into the output, we don't get a mix of* H *and* V *out the input. In terms of quantum mechanics, this device can be used to measure the polarization state of a photon in the* H, V *basis, which also means it can be used to reduce the entropy of the polarization state of the light if the input state is mixed.*

1.3 Lossy elements and the Generalized Jones Vector

Most of the optical elements that you will be describing with this formalism have Jones matrices that are unitary ($\mathbf{T}^{\dagger} = \mathbf{T}^{-1}$). However, polarization-sensitive optical elements that do not necessarily transmit 100% of the light will have non-unitary Jones matrices (see, *e.g.* Ex. 1.1). When a non-unitary matrix multiplies a unit vector, the result is no longer guaranteed to be a unit vector, and it is often therefore more useful to use what we will call the *generalized Jones vector* $\vec{\psi}$. In the spirit of Eq. 1 and 2, we could define our generalized Jones vector for linear polarization to have as its entries the electric field amplitudes for the H and V polarization components

$$\vec{\psi} = \begin{pmatrix} \mathcal{E}_{\mathrm{H}} \\ \mathcal{E}_{\mathrm{V}} \end{pmatrix}. \tag{9}$$

Later, we will be extending the Jones vector formalism to allow for complex entries, but for now I will simply state that the intensity of a beam with a generalized Jones vector $\vec{\psi}$ (using the definition (9)) is given by

$$I = \frac{1}{2} c\epsilon_0 \vec{\psi}^{\dagger} \vec{\psi}. \tag{10}$$

A Jones vector $\hat{\psi}$ can be constructed from a generalized Jones vector $\vec{\psi}$ by normalizing it according to this same rule: $\hat{\psi} = \vec{\psi}/(\vec{\psi}^{\dagger}\vec{\psi})$.

1.4 Rotation of optical elements

Another payoff of the Jones calculus is that it is simple to compute the new Jones matrix of an optical element that has been rotated about the optical axis (Fig. 3(b)). One simply needs to use the standard 2×2 Cartesian rotation matrices to transform the Jones matrix. The rotation matrix for a rotation about $+z$ through angle α is given by

$$\mathbf{R}(\alpha) \equiv \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \tag{11}$$

For instance, if an optical system has a Jones matrix $\mathbf{T}$ and it is then rotated about $+z$ through an angle α, the new Jones matrix is given by

$$\mathbf{T}_{\alpha} = \mathbf{R}(\alpha)\,\mathbf{T}\,\mathbf{R}^{-1}(\alpha) \tag{12}$$

where $\mathbf{R}^{-1}(\alpha) = \mathbf{R}(-\alpha)$.

Example 1.2 *Let's find the polarization of light that can emerge from a horizontal polarizer that has been rotated by $\alpha = 45°$ about the optical axis, as shown in Fig. 3(b). We have*

$$\begin{aligned}
\mathbf{T}_{45^\circ} &= \mathbf{R}(\tfrac{\pi}{4}) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{R}(-\tfrac{\pi}{4}) \\
&= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \\
&= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},
\end{aligned}$$

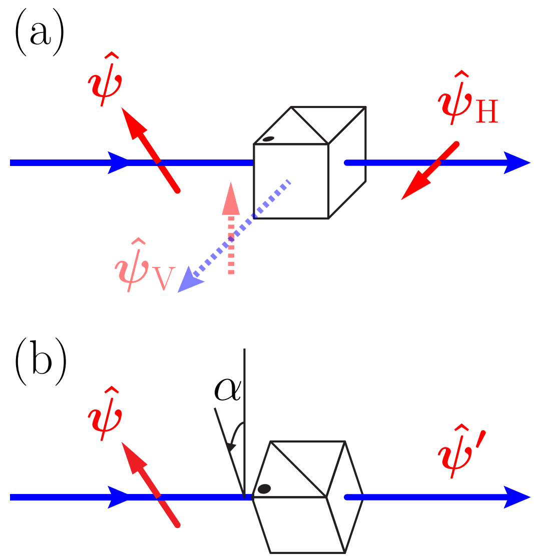

Figure 3: (a) A polarizing beamsplitter cube (PBS) can be oriented so that it only transmits horizontally-polarized light. (b) If the PBS is rotated about $+z$ through an angle α, the polarization (and power) of the output light will in general be affected.

so a general input polarization state (I will use the generalized *Jones vector* $\vec{\psi}$ *since polarizers tend to get rid of some of the input light) will lead to an output given by*

$$\mathbf{T}_{45^\circ}\vec{\psi} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \mathcal{E}_{\mathrm{H}} \\ \mathcal{E}_{\mathrm{V}} \end{pmatrix} = \left(\frac{\mathcal{E}_{\mathrm{H}} + \mathcal{E}_{\mathrm{V}}}{\sqrt{2}}\right)\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{13}$$

If you are anything like me, you may have trouble remembering which of the two rotation matrices needs to be the inverse one in Eq. (12). It turns out that there is an easy way to remember this if you consider the action of a rotated element on the polarization state:

$$\begin{aligned} \hat{\psi}' &= \mathbf{R}(\alpha)\,\mathbf{T}\,\mathbf{R}^{-1}(\alpha)\hat{\psi} \\ &= \mathbf{R}(\alpha) \quad \mathbf{T} \quad \mathbf{R}^{-1}(\alpha)\hat{\psi}. \end{aligned} \tag{14}$$

I have split up the action of the rotation matrices in Eq. 14 to point out this other interpretation for the same equation, which should be read starting on the far right: the *polarization state* is first rotated *backward* by α, then passes through the (un-rotated) polarizer, then is rotated *forward* by α. In this way, we see that the rotation matrix to the right of $\mathbf{T}$ has to be the *inverse* operation in order to make sure the optical element does the same thing to the light as if it were instead the thing that had been rotated. This is just a mnemonic to help you remember the ordering; feel free to ignore it if this does not resonate with you, but I thought I'd pass it along because I find it helpful in my own work.

1.5 Diagonal polarization

The linear polarization state that emerged from the rotated polarizer in Example 1.2 is important enough to discuss further. Consider the polarization obtained by rotating H-polarization (no matter how – just assume this can be done) by 45° about the $+z$ axis:

$$\hat{\boldsymbol{\psi}} = \mathbf{R}(\tfrac{\pi}{4})\hat{\boldsymbol{\psi}}_{\mathrm{H}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{15}$$

We will call this linear polarization state *diagonal* polarization (denoted by D), and the orthogonal state (which can be obtained by rotating $\hat{\boldsymbol{\psi}}_{\mathrm{H}}$ by $\alpha = -\pi/4$) will be called *anti-diagonal* (A). The associated Jones vectors are

$$\hat{\boldsymbol{\psi}}_{\mathrm{D}} \equiv \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{16}$$

$$\hat{\boldsymbol{\psi}}_{\mathrm{A}} \equiv \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{17}$$

The reasons these two are being given special status over, say, linear polarization that has been rotated by 9.3° from H are somewhat beyond our discussion at the moment. Essentially, $\hat{\boldsymbol{\psi}}_{\mathrm{D}}$ and $\hat{\boldsymbol{\psi}}_{\mathrm{A}}$ are eigenvectors of the Pauli matrix σ_X and therefore constitute a mathematically convenient set to use with $\hat{\boldsymbol{\psi}}_{\mathrm{H}}$ and $\hat{\boldsymbol{\psi}}_{\mathrm{V}}$ (which are eigenvectors of σ_Z) and the two senses of circular polarization, which we talk about in the next section.

1.6 Phase shifts between H and V

We have by this point fully-specified the machinery necessary for working with any pure linear polarization state. Using the basis choice of Eq. (4), if the elements of the Jones vector for a polarization state are real, the polarization is linear.

However, since $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are orthogonal, the H and V components of the light sort of "live a life of their own" in the sense that they don't depend upon one another. One thing this means is that it should be possible to shift the phase of the H and V components independently. If we create a phase difference (δ) between H and V,

$$\mathbf{E} = \tfrac{\mathcal{E}_0}{\sqrt{2}}\,\hat{\mathbf{x}}\,\mathrm{e}^{i(kz-\omega t)} + \tfrac{\mathcal{E}_0}{\sqrt{2}}\,\hat{\mathbf{y}}\,\mathrm{e}^{i(kz-\omega t-\delta)}, \tag{18}$$

we can account for this in the Jones vector if we allow it to become complex:

$$\hat{\boldsymbol{\psi}} = \begin{pmatrix} a_{\mathrm{H}} \\ |a_{\mathrm{V}}|\mathrm{e}^{-i\delta} \end{pmatrix}. \tag{19}$$

By convention, we will keep a_{H} real and positive and put the phase difference into the V component, which we can always do because Jones vectors that differ only by a global multiplicative factor are indistinguishable.[1]

[1] The only exception to this is when interference between different parts of a multipartite state become important, in which case the phase shifts are not truly "global" and must be retained.

With this generalization in hand, we can define the inner product of two Jones vectors $\hat{\boldsymbol{\psi}}_i$ and $\hat{\boldsymbol{\psi}}_j$ to be $\hat{\boldsymbol{\psi}}_i^\dagger \hat{\boldsymbol{\psi}}_j$, and the normalization condition is perhaps more natural when written as simply

$$\hat{\boldsymbol{\psi}}^\dagger \hat{\boldsymbol{\psi}} = 1. \tag{20}$$

1.7 Circular polarization

If the relative phase shift between equal-amplitude H and V components of some light is $\pm\frac{\pi}{2}$, the Jones vectors describing these states will be given by

$$\hat{\boldsymbol{\psi}}_+ \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \tag{21}$$

$$\hat{\boldsymbol{\psi}}_- \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}. \tag{22}$$

But what does this mean, for the amplitude of the V component to be the imaginary number i? We can figure this out by examining the time-dependence of the electric field more closely.

Consider the electric field associated with light described by the Jones vector $\hat{\boldsymbol{\psi}}_+$:

$$\begin{aligned}
\mathbf{E}_+ \times \tfrac{\sqrt{2}}{\mathcal{E}_0} &= \mathbb{R}\Big\{\hat{\mathbf{x}} e^{i(kz-\omega t)} + i\,\hat{\mathbf{y}} e^{i(kz-\omega t)}\Big\} \\
&= \mathbb{R}\Big\{\hat{\mathbf{x}} e^{i(kz-\omega t)} + \hat{\mathbf{y}} e^{i(kz-\omega t+\frac{\pi}{2})}\Big\} \\
&= \mathbb{R}\Big\{\hat{\mathbf{x}}\cos(kz-\omega t) + i\,\hat{\mathbf{x}}\sin(kz-\omega t) + \hat{\mathbf{y}}\cos(kz-\omega t+\tfrac{\pi}{2}) + i\,\hat{\mathbf{y}}\sin(kz-\omega t+\tfrac{\pi}{2})\Big\} \\
&= \hat{\mathbf{x}}\cos(kz-\omega t) + \hat{\mathbf{y}}\cos(kz-\omega t+\tfrac{\pi}{2})
\end{aligned} \tag{23}$$

The electric field is, of course, real (including the V component of it), but the direction it points is now *time-dependent*. We will visualize this in two different ways, the space-fixed and time-fixed cases, depicted in Fig. 4.

We can pursue the space-fixed interpretation by choosing $z = 0$ and asking how the instantaneous polarization changes in time at this location (the conclusions are of course identical for any choice of z). Using the coordinate system shown in Fig. 4, we can view the xy plane in its typical orientation by imagining the light is shining out of the page (*i.e.* Fig. 4(a) and (c). Note that we are looking directly into the light, an activity that is not recommended with lasers if you value sight.). Plugging $z = 0$ in to Eq. (23), we have a unit vector whose orientation changes according to

$$\frac{\mathbf{E}_+(z=0,t)}{\mathcal{E}_0} = \frac{\hat{\mathbf{x}}}{\sqrt{2}}\cos(\omega t) + \frac{\hat{\mathbf{y}}}{\sqrt{2}}\cos(\omega t - \tfrac{\pi}{2}). \tag{24}$$

The y-component *lags behind* the x-component by 90°, so the unit vector rotates counter-clockwise (CCW) in the xy plane, as shown in Fig. 4(a). This rotation can be associated with a rotational velocity $\boldsymbol{\omega} = \omega\hat{\mathbf{z}}$ that points in the same direction as the propagation of the light, and such light is said to have *positive helicity*, or a helicity equal to +1 (hence the designation of this with a subscript "+").

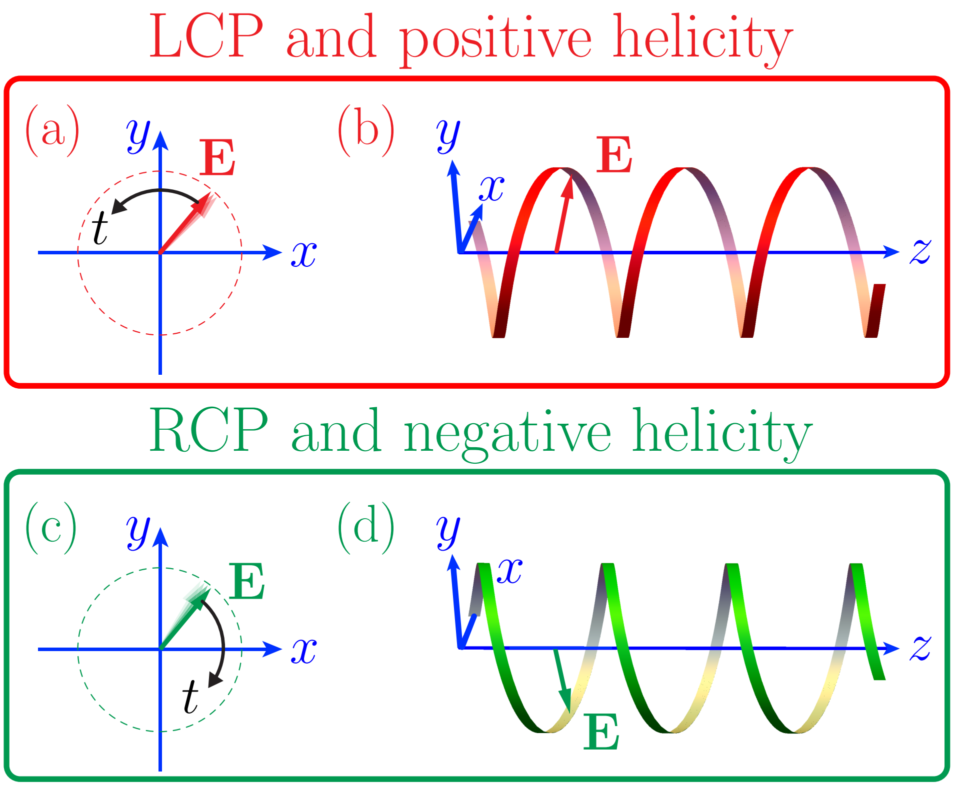

Figure 4: Space-fixed ((a) and (c)) and time-fixed ((b) and (d)) images of the electric field direction for circularly-polarized light. The naming conventions we use are shown in red and green, as well as Table 1. Note that there is no satisfactory naming convention for both the left and right pictures. We have chosen to follow the convention of Jackson [1], where the helicity is either positive or negative, and the handedness of the light matches the handedness of the spatial helix traced out by the tip of the E-field vector at any fixed instant of time.

If we had done the same exercise with $\hat{\boldsymbol{\psi}}_-$, we would have found that its space-fixed electric field rotates clockwise (CW) in the xy plane (Fig. 4(c)). Since the projection of this rotation sense on the direction of propagation is negative, this polarization state has *negative helicity*, or a helicity equal to -1.[2]

The complimentary picture we will develop is that of the time-fixed orientation of the electric field vector as a function of space. We can again choose $\hat{\boldsymbol{\psi}}_+$ and this time fix $t = 0$ to see which direction $\mathbf{E}$ points as a function of z position. Here we have

$$\frac{\mathbf{E}_+(z, t=0)}{\mathcal{E}_0} = \frac{\hat{\mathbf{x}}}{\sqrt{2}} \cos(kz) + \frac{\hat{\mathbf{y}}}{\sqrt{2}} \cos(kz + \tfrac{\pi}{2}), \tag{25}$$

[2]In the quantum treatment, the rotational velocity of the electric field vector points the same direction as the angular momentum of the photon. In particle physics, the helicity of a moving particle is defined to be positive if the projection of its angular momentum on its linear momentum is positive, and negative if it is negative. Naturally, this definition is reference-frame dependent for massive particles.

Electric Field Direction	Jones Vector	Handedness	Helicity
$\hat{\mathbf{x}}\cos(kz-\omega t)+\hat{\mathbf{y}}\cos(kz-\omega t+\frac{\pi}{2})$	$\hat{\boldsymbol{\psi}}_+ = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}$	left circular (LCP)	positive (+1)
$\hat{\mathbf{x}}\cos(kz-\omega t)+\hat{\mathbf{y}}\cos(kz-\omega t-\frac{\pi}{2})$	$\hat{\boldsymbol{\psi}}_- = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}$	right circular (RCP)	negative (-1)

Table 1: Naming conventions we will adopt for light with circular (and, with obvious generalization, elliptical) polarization. The light is assumed to be propagating along $+z$, ($\mathbf{k} = \hat{\mathbf{z}}k$), as shown in Fig. 4.

which is shown in Fig. 4(b). As z increases, we see that the x-component lags behind the y-component, and the electric field traces out a *left-handed helix* in space. Following the convention of Jackson [1], light that traces out a time-fixed left-handed helix is called *left circular polarized* (LCP) or "left-hand circular polarized." Likewise, you should be able to convince yourself that the $\hat{\boldsymbol{\psi}}_-$ polarization traces out a right-handed helix (Fig. 4(d)), and is therefore called *right circular polarized* (RCP).

It turns out that there is a one-to-one correspondence between the handedness and the helicity of circularly-polarized light, which are summarized in Table 1. So really, we did not need to separately define each one, so why have I belabored this point? My motivation for going through all of it in such detail was to define as clearly as possible the conventions we use in this course (a right-handed helix means right circular polarized, which has negative helicity, and vice versa).

That is because there are competing conventions! Many people choose to define RCP as the light that traces out a left-handed helix in space. But this is not as silly as that sentence makes it seem; the rotation sense of the space-fixed electric field vector for this light points along the propagation direction according to the right-hand-rule. I'll claim that I have never seen an elegant and intuitive solution to this problem, and it is in the end such a mess that we have just have to pick one and I have gone with Jackson [1] for our definition.

Last, though we haven't discussed elliptical polarization yet, it is worth mentioning that the same convention applies when deciding upon the handedness and helicity: the handedness will match that of the spatial helix (now a sort of squashed one) and the helicity will match the projection of the space-fixed electric field vector rotation sense on the propagation direction.

Example 1.3 *Consider a beam of RCP light that propagates toward $+z$ until it encounters a perfect mirror at normal incidence. What is the helicity and handedness of the reflected light?*

One way to figure this out is to imagine standing behind the mirror and looking through it (i.e. looking in the negative z-direction), from which viewpoint you would see that the electric field vector at the surface of the mirror must rotate CW as a function of time for both the incident and reflected light. (If the mirror were, for instance, a perfect conductor, the reflected wave must exactly cancel the incident one at all times at the surface so that the *tangential component of the electric field can continuously transition into the bulk, where it is zero.)*

Mathematically, if we retain the original coordinate system even after the reflection, the reflection is simply effected (to within a global minus sign) by the replacement $k \to -k$ where k is just some positive number with dimensions of inverse length. The spatial helix traced out by this new wave at some instant of time (call it $t = 0$) is now left-handed*:*

$$\frac{\mathbf{E}_{\mathrm{r}}(z, t=0)}{\mathcal{E}_0} = \frac{\hat{\mathbf{x}}}{\sqrt{2}} \cos(kz) + \frac{\hat{\mathbf{y}}}{\sqrt{2}} \cos(kz + \tfrac{\pi}{2}). \tag{26}$$

Alternatively if we instead wanted to express the reflected wave in a new (right-handed) coordinate system (x', y', z') so that $+z'$ is the direction it propagates, x' is horizontal, and y' is vertical, we need the replacements $(k \to -k, x \to -x', y \to y', z \to -z')$:

$$\frac{\mathbf{E}_{\mathrm{r}}(z', t)}{\mathcal{E}_0} = \frac{-\hat{\mathbf{x}}'}{\sqrt{2}} \cos(kz' - \omega t) + \frac{\hat{\mathbf{y}}'}{\sqrt{2}} \cos(kz' - \omega t - \tfrac{\pi}{2}). \tag{27}$$

This is perhaps easiest to visualize in the space-fixed picture (as a function of time), where the E-field vector rotates CCW in the $x'y'$ plane (when viewed looking from $+z'$ toward $-z'$, as usual).

So a normal-incidence reflection turns RCP into LCP, and I will state without proving it that the converse is also true. Furthermore, this holds even for reflections that happen at finite angles of incidence, though for a non-ideal (i.e. realistic) mirror, the Fresnel coefficients will typically not perfectly reflect both linear polarizations. For real mirrors, the reflected light can therefore emerge elliptically polarized, and in the case of a reflection at an angle of incidence beyond Brewster's angle, the elliptically-polarized light that reflects can even retain the same handedness as the incident circular polarized light. However, for a purely conceptual, ideal mirror ($r_\perp = r_\parallel \equiv 1$), this whole business is in keeping with the general physics notion of chirality*, a property possessed by any object that is not identical to its mirror image (such as a helix or a 3D coordinate system).*

1.8 Elliptical polarization

Now that we have a handle on linear and circular polarization, we are in a position to note that these are just special cases of the most general pure polarization state for monochromatic light, which is *elliptical polarization*. The most general pure polarization is described by the Jones vector

$$\hat{\boldsymbol{\psi}} \equiv \begin{pmatrix} a_{\mathrm{H}} \\ a_{\mathrm{V}} \mathrm{e}^{-i\delta} \end{pmatrix} \tag{28}$$

with a_{H}, a_{V}, and δ real and subject to the normalization condition $a_{\mathrm{H}}^2 + a_{\mathrm{V}}^2 = 1$.

For elliptically polarized light, if we observe the electric field vector at a fixed location in space, it will trace out an ellipse in time, such as the ellipse shown in Fig. 5. We will use the convention that this ellipse is normalized under the definition $c^2 + d^2 \equiv 1$, and we can define an *ellipticity*

$$r \equiv \frac{d}{c}. \tag{29}$$

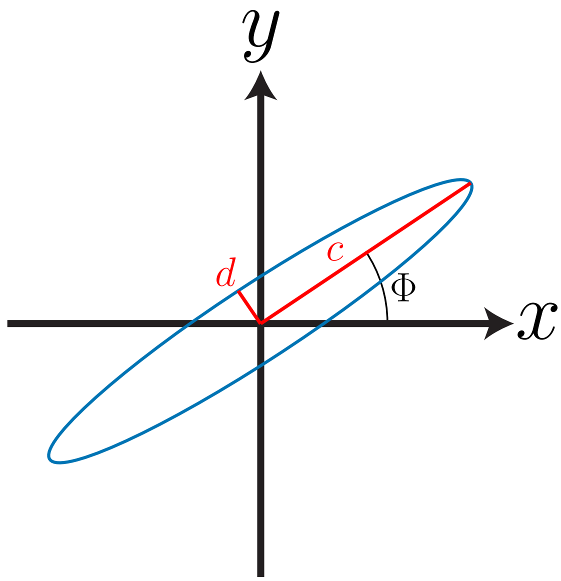

Figure 5: An ellipse with its semi-major axis, semi-minor axis, and inclination labeled by c, d, and Φ.

Clearly, the limiting cases of $r = 0$ and $r = 1$ correspond to linear and circular polarizations, respectively.

We can use some basic trigonometry to relate r and Φ to a_{H}, a_{V}, and δ via

$$a_{\mathrm{H}} = \frac{1}{\sqrt{2}}\sqrt{1-\left(\frac{r^2-1}{r^2+1}\right)\cos(2\Phi)} \tag{30}$$

$$a_{\mathrm{V}} = \frac{1}{\sqrt{2}}\sqrt{1+\left(\frac{r^2-1}{r^2+1}\right)\cos(2\Phi)} \tag{31}$$

$$\delta = -\mathrm{sign}\left[\sin(2\Phi)\right]\arccos\left(\frac{r^2-1}{\sqrt{1+r^4+r^2\cot^2(\Phi)+r^2\tan^2(\Phi)}}\right). \tag{32}$$

1.9 Introduction to Waveplates and Birefringence

Consider an optical element that could introduce a phase lag of δ on, say, only the V component of an input state:

$$\mathbf{T}(\delta)\begin{pmatrix} a_{\mathrm{H}} \\ a_{\mathrm{V}}\mathrm{e}^{i\phi_o} \end{pmatrix} = \begin{pmatrix} a_{\mathrm{H}} \\ a_{\mathrm{V}}\mathrm{e}^{i(\phi_o-\delta)} \end{pmatrix}. \tag{33}$$

The Jones matrix that can do this is given by

$$\mathbf{T}(\delta) = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{-i\delta} \end{pmatrix}. \tag{34}$$

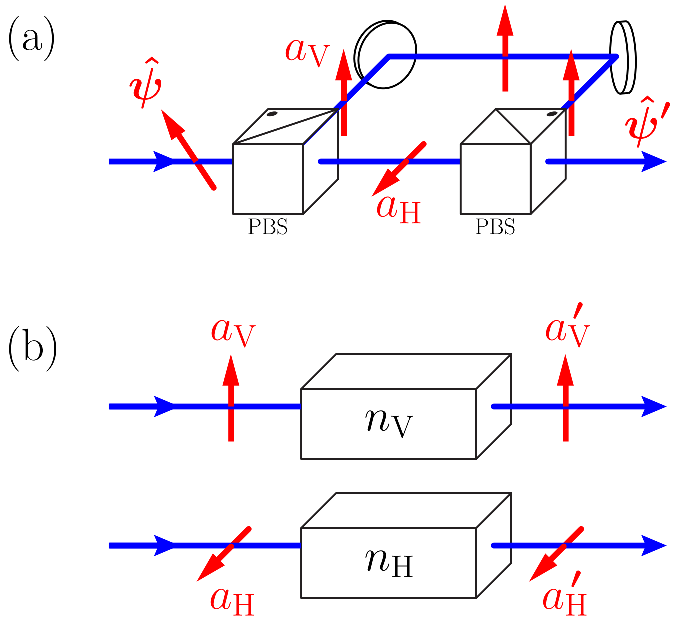

Figure 6: Optical elements that can introduce a phase difference between the H and V components of incident light. A conceptual design for a tunable waveplate is shown in (a): two polarizing beamsplitter (PBS) cubes can be used to make an imbalanced Mach-Zehnder interferometer that introduces a tunable delay for the V component. Of course, such a device would be sensitive to wavelength-scale displacements and vibrations, so this is not a practical instrument. Birefringent materials, on the other hand (shown in (b)), can have different indices of refraction (*i.e.* different phase velocities) for the H and V polarized components of light, and are more robust. The index difference will introduce a phase shift between the two polarizations that is proportional to the thickness of the material. By carefully choosing the proper material thickness, elements such as quarter-wave and half-wave plates can be made.

This type of optical element is generically called a waveplate, and before we get into the physics of how these work, I will state that there are two special cases of waveplates that are very common in optics, *half-wave plates* ($\delta = \pi$, often denoted by "HWP" or "$\lambda/2$") and *quarter-wave plates* ($\delta = \pi/2$, often denoted by "QWP" or "$\lambda/4$").

Conceptually, a waveplate can be made from optics parts made of isotropic materials by setting up a polarization-sensitive, imbalanced Mach-Zehnder interferometer such as that shown in Fig. 6(a). By tuning the delay of the V-polarized component of the light relative to the H-component, the output state will be described by Eq. (33). However, if this phase delay needs to be stable to a small fraction of a radian, the relative path lengths of the two arms of the interferometer need to be stable to a fraction of an optical wavelength, which is why this conceptual design would be unlikely to be used frequently in practice.

Much simpler is to employ crystalline materials that exhibit *birefringence*, a polarization-dependent index of refraction. For a material that exhibits linear birefringence in the H and

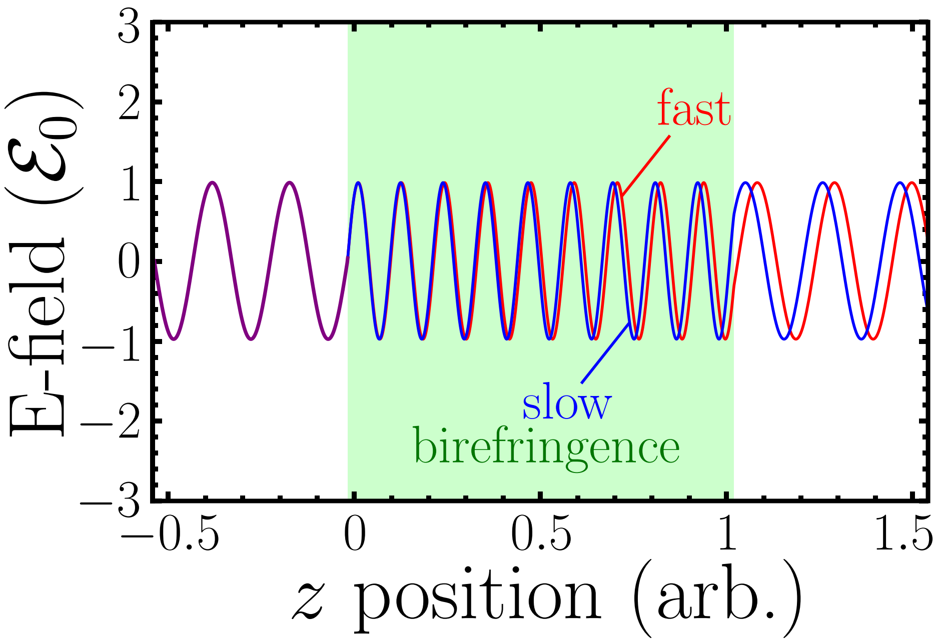

Figure 7: Fixed-time electric fields of equal amplitudes of light for polarizations aligned with the fast (red) and slow (blue) axes of a birefringent material. Between the two, the field along the slow axis has the shorter wavelength and this figure illustrates how it therefore emerges with a spatial phase *lag*.

V directions such as shown in Fig. 6(b), since the phase velocity of light decreases with increasing refractive index n, the axis that has the lower index of refraction is called the "fast axis," and the one with the larger value of n is called the "slow axis." This can be seen by noting that the temporal frequency of the light is the same inside and outside the material, which allows us to relate the refractive index to the wavelength in the material, λ' ($\equiv 2\pi/k'$):

$$\lambda' \frac{\omega}{2\pi} = \frac{c}{n}. \tag{35}$$

So light polarized along the fast axis has a *longer* wavelength and *smaller* wavenumber (k') than light polarized along the slow axis, as shown in Fig. 7. This means that when all of the light emerges from the birefringent material, $k'_{\text{fast}} z < k'_{\text{slow}} z$, and the light polarized along the fast axis has a slightly smaller (in the sense that one purely imaginary number may be smaller than another) exponent in the $\exp(ikz - i\omega t)$ term, which means it as accumulated less phase.

If we examine the time evolution of the two waves at a fixed position (say, at the exit of the birefringent crystal), this shows up as a temporal phase *delay* for the fast-axis light relative to the slow-axis light, since the slow-axis light had a larger number of oscillations while in the birefringent material. If, instead, we are interested in the spatial locations of, say, the electric field maxima at a fixed instant of time, this effect shows up as a spatial phase *advance* for the fast-axis light relative to the slow-axis light.

Once the two polarizations have exited the birefringent material, they again have the same wavelength (same k) and the difference can be summed up as a phase shift δ for the fast-axis polarization:

$$\begin{aligned} E_{\text{slow}} &\propto \mathrm{e}^{i(kz-\omega t)} \\ E_{\text{fast}} &\propto \mathrm{e}^{i(kz-\omega t-\delta)}. \end{aligned} \tag{36}$$

The transformation described by the Jones matrix in Eq. (34) can therefore be produced by a piece of linearly-birefringent material with its fast axis aligned along V.

1.10 The Half-Wave Plate

As was mentioned briefly above, a wave plate that introduces a relative phase delay of $\delta = \pi$ between two, orthogonal linear polarization components is called a half-wave plate. If the fast axis is oriented to coincide with the vertical polarization axis, its Jones matrix is given by[3]

$$\mathbf{T}_{\lambda/2}^{(\text{fast}\|\text{V})} = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{-i\pi} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{37}$$

Half-wave plates are useful because they can be used to rotate the polarization of incident light by an amount that depends on the orientation of the waveplate itself. It is therefore

[3] Note that this is the same Jones matrix we would obtain were the fast axis to be oriented in the horizontal direction modulo an overall phase shift.

most common to see HWPs mounted in rotation mounts that allow the user to controllably rotate the polarization of a laser beam passing through.

To see this mathematically, consider a HWP that is positioned with its fast axis vertical. The output polarization for an arbitrary input state will be given by

$$\hat{\boldsymbol{\psi}}' = \mathbf{T}_{\lambda/2}^{(\text{fast}\|\text{V})} \begin{pmatrix} a_{\text{H}} \\ a_{\text{V}} \end{pmatrix} = \begin{pmatrix} a_{\text{H}} \\ -a_{\text{V}} \end{pmatrix}. \tag{38}$$

Now consider the effect of the same wave plate that has been rotated about the $+z$ axis through an angle α. The output polarization is given by

$$\begin{aligned} \hat{\boldsymbol{\psi}}'' &= \mathbf{T}_{\lambda/2}(\alpha)\hat{\boldsymbol{\psi}} \\ &= \mathbf{R}(\alpha)\,\mathbf{T}_{\lambda/2}^{(\text{fast}\|\text{V})}\,\mathbf{R}(-\alpha)\hat{\boldsymbol{\psi}} \\ &= \begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix} \begin{pmatrix} a_{\text{H}} \\ a_{\text{V}} \end{pmatrix} \\ &= \begin{pmatrix} a_{\text{H}}\cos(2\alpha) + a_{\text{V}}\sin(2\alpha) \\ a_{\text{H}}\sin(2\alpha) - a_{\text{V}}\cos(2\alpha) \end{pmatrix}. \end{aligned} \tag{39}$$

Looking closely at Eq. (39), we see that it is related to Eq. (38) by a rotation through an angle $\theta \equiv 2\alpha$:

$$\hat{\boldsymbol{\psi}}'' = \mathbf{R}(2\alpha)\hat{\boldsymbol{\psi}}'. \tag{40}$$

Therefore, as one rotates a $\lambda/2$ plate, it rotates the polarization of linearly-polarized light passing through it by some constant amount plus twice the angle it has been turned. Specifically, if the angle between the incoming light's linear polarization plane and the slow axis of a $\lambda/2$ plate is γ [4], the half-wave plate rotates the light's polarization axis by 2γ about the $+z$ axis. This will also apply to the orientation of the semi-major and semi-minor axes of elliptical polarization, which we discuss later. Note that when circular polarization passes through a half-wave plate (no matter its orientation), its handedness is reversed, *viz.*

$$\begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos(2\alpha) + i\sin(2\alpha) \\ \sin(2\alpha) - i\cos(2\alpha) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \tag{41}$$

where we have neglected the global phase shift term in writing the final polarization state above.

1.11 The Quarter-Wave Plate

If the phase shift a waveplate introduces between H and V is $\delta = \pi/2$, it is known as a quarter-wave plate (QWP, or simply "$\lambda/4$," which is typically pronounced "lambda by four").

[4] By this it is meant that the polarization would need to be rotated through an acute angle γ about the $+z$ axis to make it coincide with the HWP slow axis. It is therefore possible that γ is negative.

One of the useful properties of quarter-wave plates is that they can change linearly-polarized light into circularly-polarized light, and vice versa. For this task, the input polarization axis must be aligned 45° from the fast/slow axes of the QWP:

$$\mathbf{T}_{\lambda/4}^{(\text{fast}\|\text{V})}\hat{\boldsymbol{\psi}}_{\text{D}} = \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \hat{\boldsymbol{\psi}}_{-}. \tag{42}$$

Likewise, this same, V-oriented QWP would turn anti-diagonal light ($\hat{\boldsymbol{\psi}}_{\text{A}}$) into LCP ($\hat{\boldsymbol{\psi}}_{+}$).

If incident, linearly-polarized light is incident on a QWP whose fast axis is not necessarily aligned at 45° (or 0° or 90°) from the polarization direction, the output will be elliptically-polarized. We discuss elliptical polarization in the next section, but will write the general form for a QWP whose fast axis is inclined at an angle α from V (again neglecting global phase shifts):

$$\mathbf{T}_{\lambda/4}(\alpha) = \mathbf{R}(\alpha)\,\mathbf{T}_{\lambda/4}^{(\text{fast}\|\text{V})}\,\mathbf{R}(-\alpha) = \begin{pmatrix} \cos^2(\alpha) - i\sin^2(\alpha) & (1+i)\cos(\alpha)\sin(\alpha) \\ (1+i)\cos(\alpha)\sin(\alpha) & \sin^2(\alpha) - i\cos^2(\alpha) \end{pmatrix} \tag{43}$$

1.12 Creation of arbitrary, pure polarization states in the lab

Next, we will discuss how to create any desired pure polarization state (with specific r and Φ) from an initially H-polarized laser beam[5]. The procedure will be broken into three steps:

I. Using a $\lambda/2$ plate, create the correct ratio between a_{H} and a_{V}.

II. Send this light through a $\lambda/4$ plate to make it elliptically polarized, which sets r.

III. Rotate both elements so that the elliptical output polarization has the correct inclination (Φ) in the lab.

We will now cover each of these in more detail.

I. Send H light through a HWP to set $\frac{a_{\text{V}}}{a_{\text{H}}} = r$

We have already discussed how to construct the Jones matrix for an HWP whose slow axis has been rotated by an angle θ (about $+z$) from the horizontal,

$$\mathbf{T}_{\lambda/2}(\theta) = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix}. \tag{44}$$

If we operate on $\hat{\boldsymbol{\psi}}_{\text{H}}$ with this, we find that the output polarization is

$$\hat{\boldsymbol{\psi}}' = \begin{pmatrix} \cos(2\theta) \\ \sin(2\theta) \end{pmatrix}. \tag{45}$$

[5]We will ignore the handedness of the desired state for this exercise, but it is straightforward to flip this if the one generated by this procedure is incorrect (see Eq. (41)).

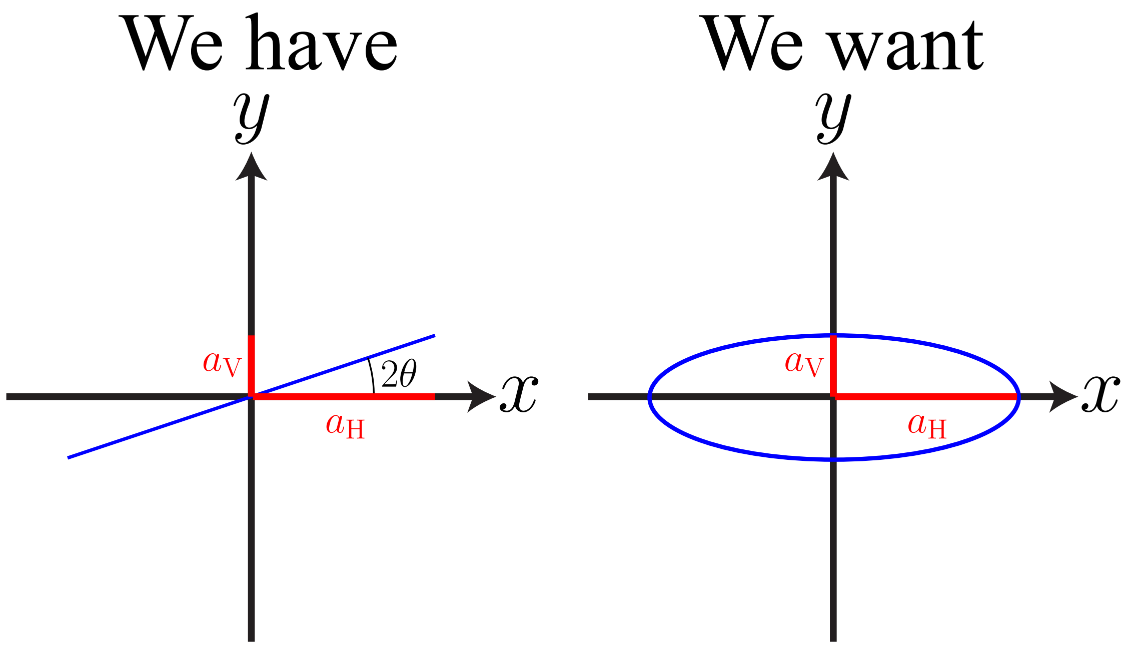

Figure 8: The polarization state created in step I has $a_\mathrm{V}/a_\mathrm{H}$ equal to the desired ellipticity, r. Shown here is the case where $\theta \leq \pi/8$, on the left. If a relative phase shift of $\pi/2$ is introduced between the two components, this creates elliptical polarization with ellipticity r, shown on the right.

Since we require $\frac{a_\mathrm{H}}{a_\mathrm{V}} = r^{-1}$, we have

$$r^{-1} = \frac{\cos(2\theta)}{\sin(2\theta)} = \frac{1}{\tan(2\theta)} \tag{46}$$

and we get an expression for θ,

$$\theta = \frac{1}{2}\arctan(r). \tag{47}$$

II. Use a properly-oriented QWP to convert this to elliptical polarization

After the first HWP, we still have linear polarization, which you can verify by noting that both entries in the Jones vector are real in Eq. 45. We now want to introduce a relative phase of $\frac{\pi}{2}$ between the H and V components, which will create a state with the desired ellipticity, shown in Fig. 8.

We therefore see that a QWP with its slow axis along H will introduce this phase shift:

$$\hat{\boldsymbol{\psi}}'' = \mathbf{T}_{\lambda/4}^{(\text{fast}\|V)}\hat{\boldsymbol{\psi}}' = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{-i\frac{\pi}{2}} \end{pmatrix} \begin{pmatrix} \cos(2\theta) \\ \sin(2\theta) \end{pmatrix} = \begin{pmatrix} \cos(2\theta) \\ \mathrm{e}^{-i\frac{\pi}{2}}\sin(2\theta) \end{pmatrix}, \tag{48}$$

which is elliptically polarized light with the desired ellipticity, though its semi-major axis is not yet at the correct angle Φ.

III. Rotate both wave plates to tilt the polarization to the desired angle in the lab

Last, we need to rotate this whole ellipse. Imagine for a moment that we rotate our first wave plate by an additional angle $\Phi/2$ (i.e. so that it's now oriented at $\theta + \Phi/2$). If we then rotate the QWP by Φ, the result will be the same as step II (Fig. 8, right), but with the whole output tilted by Φ. This will produce the desired elliptical polarization.

References

[1] John David Jackson. *Classical Electrodynamics*. John Wiley and Sons, Inc., 3rd edition, 1999.