4. Theory

When an electric field is applied to a crystal or liquid, the refractive index of the crystal or liquid is altered. This phenomenon is called the electro-optic effect, which has been applied to many important areas in science and technology. With a short response time (10^{10} Hz), electro-optic effect can be used by shutters in high-speed photography and optical choppers in speed of light measurement. After the emergence of lasers, the research and development of electro-optic effect has advanced rapidly. So far, electro-optic devices have been widely applied to areas such as laser communication, laser ranging, laser display, and optical data-processing.

4.1 Linear electro-optic effect and refractive index ellipsoid

Typically, the change in the refractive index of an electro-optic crystal to an electric field can be expressed as

$$n = n_0 + aE_0 + bE_0^2 + ... \tag{1}$$

where a and b are constants, and n_0 is the refractive index of the crystal in the absence of an electric field. The second term in equation (1) is linearly proportional to the electric field, called the linear electro-optic effect or the Pockels effect; while the third term is the quadratic electro-optic effect, also called the Kerr effect. Pockels effect exists only in crystals that are lack of inversion symmetry while Kerr effect can exist in any materials but it is generally much weaker than the Pockels effect.

As shown in Figure 1, when light propagates in an anisotropic crystal, the propagating velocities of light in the crystal are different for different propagating directions or different oscillation directions of the electric vector (polarization), resulting in different refractive indices. Normally, a refractive index ellipsoid is used to describe the relationship among the refractive index of the material, the propagating direction and the polarization state of light.

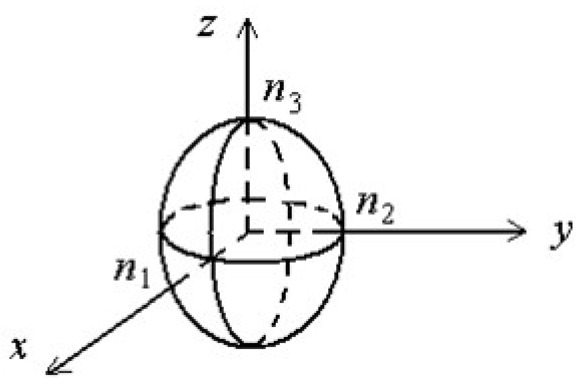

Figure 1 Refractive index ellipsoid in the absence of an electric field

In the absence of an electric field, the refractive index ellipsoid of an anisotropic crystal can be expressed with principal axis coordinates

$$\frac{x^2}{n_1^2} + \frac{y^2}{n_2^2} + \frac{z^2}{n_3^2} = 1 \tag{2}$$

where n_1, n_2, and n_3 are the principal refractive indices of the crystal along the principal axes of x, y, and z, respectively. When an electric field is applied, the shape and orientation of the above refractive index ellipsoid are subject to change and hence the ellipsoid equation becomes

$$\frac{x^2}{n_{11}^2}+\frac{y^2}{n_{22}^2}+\frac{z^2}{n_{33}^2}+\frac{2yz}{n_{23}^2}+\frac{2xz}{n_{13}^2}+\frac{2xy}{n_{12}^2}=1 \tag{3}$$

Generally speaking, the linear electro-optic effect of a crystal is classified into two categories, longitudinal electro-optic effect and transverse electro-optic effect. The former refers to as the electro-optic effect of a crystal when an electric field is applied along the propagating direction of light in the crystal; while the latter is the electro-optic effect when the electric field is applied perpendicularly to the propagating direction of light in the crystal. Normally, longitudinal electro-optic effect is used for KDP-type crystals while transverse electro-optic effect is used for $LiNbO_3$-type crystals. In this experiment, the linear electro-optic effect of a $LiNbO_3$ crystal is studied. By using the transverse electro-optic modulation technique, the half-wave voltage and electro-optic efficient of a $LiNbO_3$ crystal can be measured, and its output optical characteristic can be observed.

A $LiNbO_3$ crystal consists of a trigonal lattice (3M point group), with a third axis of rotation along the principal z-axis (the propagating direction of light in the crystal). $LiNbO_3$ is an uniaxial crystal whose refractive index ellipsoid is a rotational ellipsoid described by

$$\frac{x^2+y^2}{n_o^2}+\frac{z^2}{n_e^2}=1 \tag{4}$$

where n_o and n_e are the refractive indices of the ordinary ray and extraordinary ray, respectively. When an electric field is applied along the principal x-axis, the uniaxial crystal becomes a biaxial crystal, and the cross-sectional projection of the refractive index ellipsoid is changed from circle to ellipse, described by

$$\left(\frac{1}{n_0^2}-\gamma_{22}E_x\right)x^2+\left(\frac{1}{n_0^2}+\gamma_{22}E_x\right)y^2-2\gamma_{22}E_x xy=1 \tag{5}$$

where γ_{22} is the electro-optic coefficient of the crystal.

By taking principal axis transform, equation (5) can be rewritten as

$$\left(\frac{1}{n_0^2}-\gamma_{22}E_x\right)x'^2+\left(\frac{1}{n_0^2}+\gamma_{22}E_x\right)y'^2=1 \tag{6}$$

Considering $(n_0)^2\gamma_{22}E_x<<1$, equation (6) can be simplified as

$$n_{x'}=n_0+\frac{1}{2}n_0^3\gamma_{22}E_x \tag{7}$$

$$n_{y'}=n_0-\frac{1}{2}n_0^3\gamma_{22}E_x \tag{8}$$

Hence, equation (6) can be rewritten as

$$\frac{x'^2}{n_{x'}^2} + \frac{y'^2}{n_{y'}^2} = 1 \tag{9}$$

where *x*' and *y*' are called the inductive axes of the $LiNbO_3$ crystal.

4.2 Theory of electro-optic modulation

If a laser is used as a tool to carry information, one must load the signal to be carried onto the laser radiation. This process is called laser modulation and the apparatus used in this process is called laser modulator. On the contrary, the restoration process of carried signal from modulated laser radiation is called demodulation. Because the laser acts with a signal carrying function, it is called a carrier. Similar to radio modulation, laser modulation can take forms of continuous amplitude modulation, continuous frequency modulation, continuous phase modulation, or pulsed modulation. By comparison, intensity modulation is the most common method, in which the strength of a modulation signal is proportional to the intensity of an optical carrier so that the intensity of the laser is regulated by the modulation signal. The reason that intensity modulation is so common is because photo-receivers respond to the change in optical intensity rather than amplitude.

There are many mechanisms for laser modulation, such as mechanical modulation, electro-optic modulation, acousto-optic modulation, magneto-optic modulation, or source current modulation. Among them, electro-optic modulation is fast and simple, and hence it is widely applied to laser modulation technology and mixed optical bistable devices.

4.2.1 Transverse electro-optic modulation

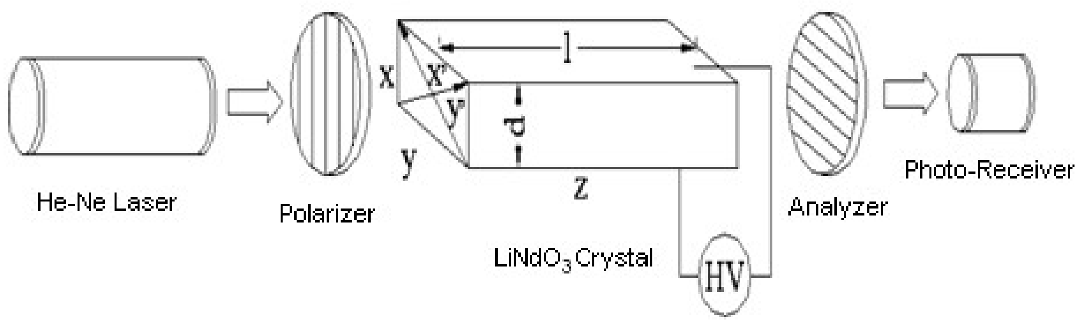

Figure 2 Schematic of transverse electro-optic modulation

Figure 2 shows the schematic of a laser amplitude modulator using the transverse electro-optic effect of a $LiNbO_3$ crystal. The polarization of a polarizer is aligned along the *x*-axis while the polarization of an analyzer is parallel to the *y*-axis. Thus, the projections of the incident light on axes *x*' and *y*' at the surface of incidence (*z*=0) are identical with complex amplitudes expressed as

$$E_{x'}(0) = A, \quad E_{y'}(0) = A \tag{10}$$

Hence, the intensity of the incident light is

$$I_i \propto \overline{E} \cdot \overline{E^*} = |E_{x'}(0)|^2 + |E_{y'}(0)|^2 = 2A^2 \tag{11}$$

After going through an electro-optic crystal with a length of L, a phase delay of δ is produced between the two components

$$E_{x'}(L) = A, \quad E_{y'}(L) = Ae^{-i\delta} \tag{12}$$

After transmitting through the analyzer, the resultant complex amplitude is the addition of the x' and y' components projected on the y-axis

$$\left(E_y\right)_o = \frac{A}{\sqrt{2}}\left(e^{i\delta} - 1\right) \tag{13}$$

Correspondingly, the intensity of light transmitted through the analyzer is

$$I_t \propto \left[\left(E_y\right)_o \left(E_y\right)_o^*\right] = \frac{A^2}{2}\left[\left(e^{i\delta} - 1\right)\left(e^{-i\delta} - 1\right)\right] = 2A^2 \sin^2\frac{\delta}{2} \tag{14}$$

Based on equations (11) and (14), the transmittance of the crystal is

$$T = \frac{I_t}{I_i} = \sin^2\frac{\delta}{2} \tag{15}$$

From equations (7) and (8), the phase delay can be further expressed as

$$\delta = \frac{2\pi}{\lambda}\left(n_{x'} - n_{y'}\right)L = \frac{2\pi}{\lambda} n_0^3 \gamma_{22} U \frac{L}{d} \tag{16}$$

where U is the applied voltage, d is the thickness of the crystal, and λ is the wavelength of the laser. Obviously, the phase delay depends on the voltage applied. When the applied voltage is increased to a certain value, a phase delay of π can be achieved, corresponding to an optical path difference of $\lambda/2$ and 100% transmittance. Such voltage is called the half-wave voltage of the crystal denoted by U_π, an important parameter to describe the electro-optic characteristic of an electro-optic crystal. For a given crystal, a smaller half-wave voltage is preferred, which means smaller modulation voltage can be used.

From equation (16), the half-wave voltage can be expressed as

$$U_\pi = \frac{\lambda}{2 n_0^3 \gamma_{22}}\left(\frac{d}{L}\right) \tag{17}$$

Obviously, the half-wave voltage depends on the dimensions of the crystal, linearly proportional to the length of the crystal but inversely proportional to the spacing between electrodes.

Substituting equation (17) to (16), we get

$$\delta = \pi \frac{U}{U_\pi} \tag{18}$$

Thus, equation (15) can be rewritten as

$$T = \sin^2\frac{\pi}{2U_\pi}U = \sin^2\frac{\pi}{2U_\pi}\left(U_0 + U_m \sin wt\right) \tag{19}$$

where U_0 and $U_m\sin(wt)$ are the dc and ac voltages applied to the crystal, respectively; U_m is the amplitude of the ac voltage and w is the modulation frequency. For a given crystal with a known laser wavelength, the half-wave voltage is a constant, and then the transmittance of the crystal depends solely on the applied voltage.

4.2.2 Effect of dc bias voltage on optical output characteristic of electro-optic crystal

1. If $U_0=U_\pi/2$, and $U_m<<U_\pi$, as shown in Figure 3 (a), the modulation voltage should be selected within a certain range to achieve a linear modulation with high-efficiency. Now, equation (19) can be rewritten as

$$T = \sin^2\left(\frac{\pi}{4}+\frac{\pi}{2U_\pi}U_m \sin wt\right) = \frac{1}{2}\left[1-\cos\left(\frac{\pi}{2}+\frac{\pi}{U_\pi}U_m \sin wt\right)\right] = \frac{1}{2}\left[1+\sin\left(\frac{\pi}{U_\pi}U_m \sin wt\right)\right] \quad (20)$$

Considering $U_m<<U_\pi$, equation (20) can be approximated as

$$T \approx \frac{1}{2}\left[1+\frac{\pi U_m}{U_\pi}\sin wt\right] \propto \sin wt \quad (21)$$

Although the amplitudes of the output and modulation signals are different, their frequencies are identical and therefore the output signal is distortion free, called linear modulation.

2. If $U_0=0$, and $U_m<<U_\pi$, as shown in Figure 3 (b), equation (19) becomes

$$T = \sin^2\left(\frac{\pi}{2U_\pi}U_m \sin wt\right) \approx \frac{1}{4}\left(\frac{\pi}{U_\pi}U_m\right)^2 \sin^2 wt \approx \frac{1}{8}\left(\frac{\pi}{U_\pi}U_m\right)^2 \left(1-\cos 2wt\right) \propto \cos 2wt \quad (22)$$

It is apparent from equation (22) that the output frequency is twice of the modulation frequency, resulting in frequency-doubling distortion.

3. If $U_0=U_\pi$, and $U_m<<U_\pi$, then equation (19) can be approximated as

$$T \approx 1-\frac{1}{8}\left(\frac{\pi U_m}{U\pi}\right)^2 \left(1-\cos 2wt\right) \propto \cos 2wt \quad (23)$$

Similarly, the output signal still suffers from frequency-doubling distortion.

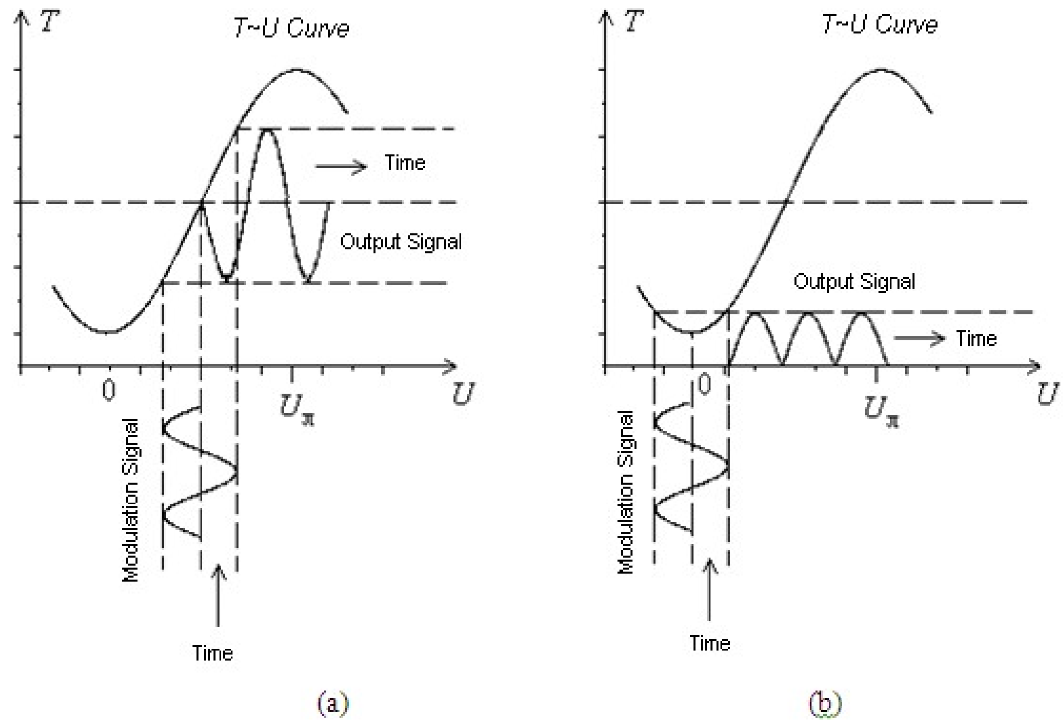

Figure 3 Effect of dc voltage bias on output characteristic of electro-optic crystal

In summary, if dc voltage bias U_0 is selected at 0 V or near the half-wave voltage U_π, nonlinear electro-optic modulation occurs, and the output signal suffers from distortion. If $U_0=U_\pi/2$, but $U_m>U_\pi$, nonlinear electro-optic modulation still occurs because the requirement of small-signal modulation is not met. In the latter case, though the dc voltage bias is selected in the linear region, the output signal still suffers from distortion.