Appendix C

Polarization Sensitive Optical Elements

Dichroism (anisotropy in absorption) and birefringence (double refraction), discussed below, are two properties of materials which have been very effectively employed to produce optical elements with specific polarization behavior. The next section gives a brief description of these two properties. Examples of two polarization sensitive optical components, which have been used in IMPOL, are used to illustrate the application of birefringence for this purpose.

C.1 Dichroism and Birefringence

The term dichroism is a misnomer and worse still it has been used, in the context of optics itself, to denote at least four different phenomena (see Shurcliff 1962, p. 43). The only property, common to all these phenomena, is that they all show some colour effects under different conditions, but basically due to different physical mechanisms. The term, as it is used here, refers to the property of materials by which they show difference in absorption of light depending on its polarization properties and often also the direction of propagation through the material.

Light propagating through a dichroic material is decomposed into two orthogonal polarization components (or eigen vectors), each of which is absorbed to a different extent by the material. Quantitatively, dichroism may be defined as the difference in absorption coefficients for the two polarization components. Dichroic materials are extensively used to produce the commercially available polarizers. It is possible to induce dichroic property in certain materials. For example, the most common types of sheet polarizers are made by processes involving unidirectional stretching of large plastic sheets such as the polymer, polyvinyl alcohol.

Birefringence or double refraction, as the name indicates, is the property by which certain materials have two different refractive indices for two orthogonal polarization

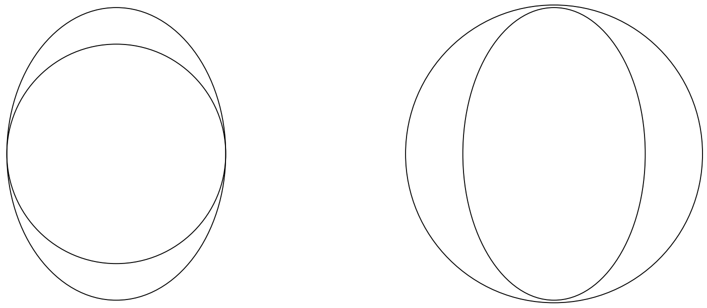

Figure C.1: Wavefronts in (a) negative ($n_\circ < n_e$) and (b) positive ($n_\circ > n_e$) uniaxial crystals

components (eigen vectors). This property is observed mostly in crystals like quartz and calcite, which have an anisotropic crystal structure, but have an axis of symmetry which is known as the principal axis. In general, such crystals can also have two preferred directions called optic axes, which are associated with their optical properties. But in general, the optic axes need not coincide with any of the axes of crystal symmetry. Some crystals are uniaxial in the sense that the two optic axes coincide.

It was suggested by Huygens (1629-1695) that the propagation of light in a uniaxial crystal could be understood by assuming that the light is split into two components. The one which propagates with a spherical wavefront and obeys both laws of refraction is called the ordinary ray. The second wavefront is an ellipsoid of revolution and is called the extraordinary ray. The fast axis of the crystal is the direction in which light travels faster and the orthogonal direction is called the slow axis. The ellipsoid touches the spheroid along the fast axis for a positive uniaxial crystal and along the slow axis for a negative uniaxial crystal (see Fig. C.1). If light is incident normally on one surface of a plate, made of birefringent uniaxial crystal, the ordinary ray will pass straight through, but the extraordinary ray, in general, will emerge parallel to the ordinary one, but shifted laterally by an amount which is proportional to the thickness of the plate.

C.1.1 Half-wave plate

The ordinary and the extraordinary rays travel with different velocities within a birefringent material and hence acquire a phase difference between them as they propagate. A plate made of such a material, with optically flat surfaces and having a thickness such that the path difference between the ordinary and the extraordinary beams corresponds to $\lambda/2$ at some wavelength, acts as a half-wave plate at the wavelength (see, for example, Serkowski 1974 for a more general discussion). Essentially, the equation governing the half-wave plate is

$$(n_{\text{o}} - n_{\text{e}}) \times t = \frac{\lambda}{2}\,, \tag{C.1}$$

where n_{o} and n_{e} are the refractive indices of the ordinary and the extraordinary rays and t is the thickness of the plate. Thus the two orthogonally polarized components (parallel and perpendicular to the fast-axis of the HWP), acquire a phase difference of π rad, when light passes through the HWP. If the incident light is plane-polarized at an angle θ with respect to the fast-axis of the HWP, since one of the polarization components acquires an additional phase difference of π rad, the emerging light will be polarized at an angle $-\theta$ with respect to the fast-axis. This property is quite useful, since allows to rotate the plane of polarization of light through a known angle. For example, if the HWP is rotated through an angle $\Delta\theta$ the emergent light will rotate through an angle $2\Delta\theta$.

As Eq. C.1 illustrates, the plate introduces a relative retardance of π, between the two components, only at a particular wavelength and hence such a simple scheme is not particularly useful. Plates which retard by π at at least two different wavelengths could be made by gluing together two plates of different birefringent materials with appropriate thicknesses. Such materials should have fairly similar refractive indices, to prevent reflection from their interface, should be hard and easy to polish to high precision and should be transparent over the wavelength range of interest. Crystalline quartz and magnesium fluoride (both positive crystals) are two such materials which are commonly used for making half-wave plates. The conditions for the composite plate to have a retardance of π, at two wavelengths λ_1 and λ_2, are

$$\begin{aligned} b_1(\lambda_1)t_1 - b_2(\lambda_1)t_2 &= \frac{\lambda_1}{2} \\ b_1(\lambda_2)t_1 - b_2(\lambda_2)t_2 &= \frac{\lambda_2}{2}\,, \end{aligned} \tag{C.2}$$

where b_1 and b_2 are the difference between ordinary and extraordinary refractive indices for the first and the second plate. A way of further improving the achromatism, proposed by Pancharatnam (1955), employs three half-wave plates glued together in such a way that the first and the last have their corresponding optic axes parallel and that of the central one at an angle of 57.2° with that of the other two.

C.1.2 Wollaston Prism

A Wollaston prism, as illustrated in Fig. C.2, is essentially made of two right angle prisms glued back to back. The optic axis of each prism is perpendicular to the incident ray, but the prisms are glued in such way that their axes are mutually orthogonal (see Fig. C.2). The normally incident ray is not deviated as it enters the first prism. However, at the interface, the ordinary ray in the first prism (with vibrations orthogonal to the optic axis) becomes an extraordinary ray in the second one and since n_{e} is less than n_{o} (for positive crystals), it deviates away from the normal. On the other hand,

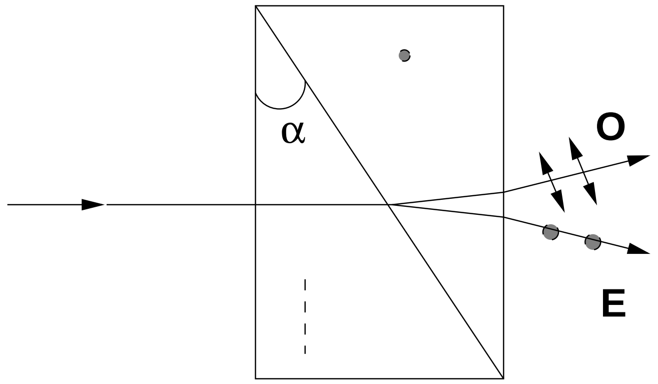

Figure C.2: Schematic of a Wollaston prism. The directions of the optic axes of the two parts are marked. The dots represent directions normal to the plane of the paper.

the extraordinary ray (with vibrations along the optic axis) in the first prism becomes an ordinary ray in the second one and refracts towards the normal at the interface. Both of these rays deviate away from the normal as they emerge from the surface of the Wollaston prism and hence their separation increases further. The total separation between the beams depends on the two refractive indices, n_e and n_o as well as the angle α.

The two rays emerge from the Wollaston prism, almost symmetrically with respect to the direction of the incident ray. Since both the rays are deviated, the Wollaston prism gives larger angular separation than many other types of polarizing beam-splitters. Although Wollaston prisms are chromatic, this becomes a serious problem only for separation angles over several degrees. One of the most common application of the Wollaston prism, as in our case, is to determine the intensities in two mutually orthogonal polarization components. Since light never travels along the optic axis in a Wollaston prism, it is not subject to optical activity and hence the intensities of the two emerging rays are directly proportional to those of the two orthogonal polarization components of the light.

C.2 Mueller matrices

A Mueller matrix $\mathbf{M}$, is a 4×4 matrix whose components define the relations between the Stokes parameter vectors $\mathbf{S}$ and $\mathbf{S}'$, representing light before and after it passes through the corresponding optical component. Thus

$$\mathbf{S}' = \mathbf{M} \cdot \mathbf{S} \ . \tag{C.3}$$

The Mueller matrix for a half-wave plate with its optic axis at a position angle α, with respect to a fiducial direction, is given by

$$\mathbf{M}_H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 4\alpha & \sin 4\alpha & 0 \\ 0 & \sin 4\alpha & -\cos 4\alpha & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} , \tag{C.4}$$

and that for an ideal analyzer with its axis at an angle θ, by

$$\mathbf{M}_p = \frac{1}{2} \begin{bmatrix} 1 & \cos 2\theta & \sin 2\theta & 0 \\ \cos 2\theta & \cos^2 2\theta & \frac{1}{2}\sin 4\theta & 0 \\ \sin 2\theta & \frac{1}{2}\sin 4\theta & \sin^2 2\theta & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} . \tag{C.5}$$

Therefore, the Mueller matrices corresponding to the ordinary and the extraordinary rays emerging from the Wollaston prism are obtained by putting $\theta = 0°$ and $\theta = 90°$ in Eq. C.5. Combining Eq. C.3 and C.4, the Stokes vector of the light emerging from the HWP is obtained in terms of that of the incident light, as

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 4\alpha & \sin 4\alpha & 0 \\ 0 & \sin 4\alpha & -\cos 4\alpha & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \begin{bmatrix} I \\ Q\cos 4\alpha + U \sin 4\alpha \\ Q \cos 4\alpha - U \sin 4\alpha \\ -V \end{bmatrix} . \tag{C.6}$$

The orthogonal polarization components produced by the Wollaston prism analyzer are given by

$$\frac{1}{2} \begin{bmatrix} 1 & \pm 1 & 0 & 0 \\ \pm 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} I \\ Q\cos 4\alpha + U \sin 4\alpha \\ Q \cos 4\alpha - U \sin 4\alpha \\ -V \end{bmatrix} = \frac{1}{2} \left[I \pm Q \cos 4\alpha \pm U \sin 4\alpha \right] \begin{bmatrix} 1 \\ \pm 1 \\ 0 \\ 0 \end{bmatrix} , \tag{C.7}$$

where the + signs are for the polarization component at $\theta = 0°$ (say, ordinary) and the − signs are for the one at $\theta = 90°$ (extraordinary). The ratio $R(\alpha)$ is obtained as

$$\begin{aligned} R(\alpha) = \frac{I_\mathrm{o} - I_\mathrm{e}}{I_\mathrm{o} + I_\mathrm{e}} &= \frac{Q \cos 4\alpha + U \sin 4\alpha}{I} \\ &= q \cos 4\alpha + u \sin 4\alpha \; . \end{aligned} \tag{C.8}$$

From Eq. A.3, we get

$$q = p \cos 2\theta \qquad \text{and} \qquad u = p \sin 2\theta \; . \tag{C.9}$$

Substituting in Eq. C.8, we get

$$R(\alpha) = p \cos\left(2\theta - 4\alpha\right) , \tag{C.10}$$

where

$$p = \sqrt{q^2 + u^2} \qquad \text{and} \qquad \theta = \frac{1}{2} \tan^{-1}\left(\frac{q}{u}\right) . \tag{C.11}$$