NR 046-708

(1)

SPECIAL REPORT 1 December 1948

THE BIREFRINGENT FILTER

ONR CONTRACT: N8onr-64801

PROJECT NO: NR-015-801

CONTRACTOR: High Altitude Observatory of Harvard University and University of Colorado, Boulder, Colorado.

PROJECT TITLE: Studies of Solar Corona

PROJECT SUPERVISORS: John W. Evans
Walter Orr Roberts

SUBMISSION DATE: 10 December 1948

39

SPECIAL NOTE:

This report summarizes the results of a long-extended series of investigations conducted by Dr. Evans at the High Altitude Observatory. His work has been done in addition to his contract activities under AMC contract W19-122 ac-17 and ONR contract N8onr-64801. It is impossible to consider this research the direct result of either contract, yet both contributed in part to the support which the High Altitude Observatory has given Dr. Evans. In addition, the results of the report are directly pertinent to both contracts, so that this report is being submitted to the full distribution lists of both contracting agencies. A title page appropriate to each contracting agency precedes this page.

In addition, this report will be distributed freely from the High Altitude Observatory with no title page. The report will also appear shortly as a published article, probably in the Journal of the Optical Society of America.

Walter Orr Roberts

WALTER ORR ROBERTS, Sup't.
High Altitude Observatory

1 December 1948

THE BIREFRINGENT FILTER

ABSTRACT

The basic principles of birefringent filter operation are briefly discussed and references given to papers which discuss the theory. Off-axis effects are investigated as well as field of view limitations, and methods for extending the field of view are considered. The split element filter is described; it uses only half as many polarizers as a conventional filter, after the first polarization. Wave length adjustment possibilities are evaluated for conventional and split element filters. The usefulness of various crystal materials is mentioned. Finally, the polarization interferometer is discussed as a way of accomplishing the narrow-band transmission of an impossibly thick birefringent filter element. Several possible forms of interferometer are mentioned.

OUTLINE

Page

I. Basic Birefringent Filter Design

A. Introduction 2

B. The Simple Birefringent Filter 2

C. Off-Axis Effects in Simple Filters 8

D. Lyot's Wide Field Filters 14

II. Recent Modifications of the Birefringent Filter

A. The Split Element Filter 18

B. Filters of Adjustable Wave Length 22

III. Materials and Applications

A. Materials for Birefringent Filters 33

B. Polarizing Interferometer Filters 34

IV. References 38

* * * *

* * * *

I. BASIC BIREFRINGENT FILTER DESIGNS

A. Introduction

During the past few years the birefringent filter has proved an effective tool in astronomical research. Its utility, however, is not confined to astronomy and the purpose of the present paper is partly to bring it to the attention of investigators in other fields.

Briefly, the birefringent filter serves the purpose of a monochromator over an extended field. It can be designed to transmit a wave length band of any desired width (down to a fraction of an angstrom) centered at any selected wave length. It is used very much like an ordinary glass or gelatin filter in either a collimated or a converging beam of light, but with some limitation in field size or focal ratio, depending on type of construction, material, and band width.

The invention of the birefringent filter is one of the many important contributions of the French astronomer, Bernard Lyot[1*], in instrumental astronomy. He first published the basic principles of its operation in 1933. Öhman[2*] independently invented the filter, and in 1938 constructed the first one to be used for solar observations, with a transmission band about 40 angstroms wide centered on the $H\alpha$ line. With it he succeeded in seeing and photographing the brighter prominences, although it was evident that a much sharper band would be necessary for the best results.

In a later paper Lyot[3*] has given a very complete discussion of the history, theory, and construction of birefringent filters. For the benefit of readers to whom his papers are not readily available, the present paper reviews enough of the elementary theory to suffice for the design of filters of any feasible characteristics. The remainder of the paper is a discussion of newer developments which serve to simplify the construction of the filters and extend their field of usefulness.

B. The Simple Birefringent Filter

Several forms of the birefringent filter are possible, differing in width of field and complexity of construction. They all depend, however, on the interference of polarized light transmitted through layers of birefringent crystal in the direction perpendicular to the plane of the optic axis, if the crystal is biaxial, or any direction perpendicular to the optic axis if the crystal is uniaxial.

Since we can regard the uniaxial crystal as a degenerate biaxial crystal, most of the following discussions will consider only the biaxial case. Let ϵ and ω be the extraordinary and ordinary indices or refraction of any uniaxial crystal and α, β, γ be the smallest, intermediate, and greatest principal indices of refraction of a biaxial crystal, respectively. Any expression for a biaxial crystal is valid for a uniaxial crystal if one of the following substitutions is made:

$$\alpha = \omega, \quad \beta = \omega, \quad \gamma = \epsilon \qquad \text{if} \quad \epsilon - \omega > 0$$

or

$$\alpha = \epsilon, \quad \beta = \omega, \quad \gamma = \omega \qquad \text{if} \quad \epsilon - \omega < 0$$

Unless otherwise specified the directions of vibration of light for which the refractive indices are α, β and γ will be referred to as the α-axis, β-axis and γ-axis.

The quantity μ is defined by

$$\mu = \gamma - \alpha$$

The term "retardation" will be used to indicate a path difference in terms of wave lengths.

For brevity, the direction of vibration of light transmitted by a polarizer (prism or film) will be referred to as the axis of the polarizer.

Consider a block of some birefringent crystal, b_1 in Figure 1, cut with its surfaces normal to its β-axis. Let light plane polarized at an angle of 45^o to the α-axis enter the crystal along the β-axis. In the crystal the light divides into two components polarized with vibrations parallel to the α and γ-axes, travelling with different velocities, $\frac{c}{\alpha}$ and $\frac{c}{\gamma}$. On emerging from the crystal the two components have therefore a relative retardation of n_1, given by:

$$n_1 = \frac{d_1}{\lambda}\mu \qquad (2.1)$$

where d_1 is the thickness of the crystal in the β direction, and λ is the wave length of the light.

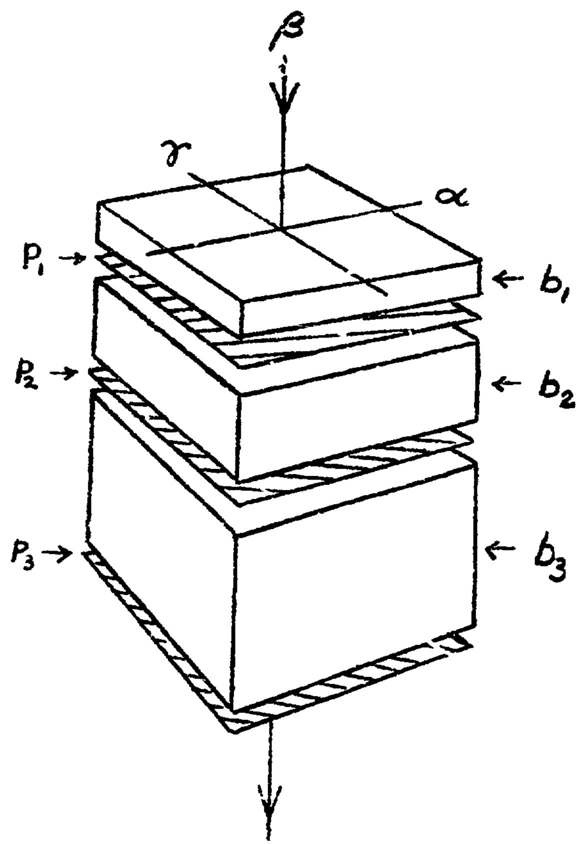


FIGURE 1. BIREFRINGENT FILTER OF THREE ELEMENTS.

If now the light traverses a polarizer, p_1, (which may be either a Nicol or similar prism, or a film polarizer) with its axis parallel to the vibration plane of the entering light, the two components interfere. The transmission, τ_1, of the b_1, p_1 combination is:

$$\tau_1 = \cos^2 \pi n_1 \qquad (2.2)$$

If white light traverses the combination, the spectrum of the emergent light consists of regularly spaced alternate bright and dark bands at wave lengths where n_1 is alternately integral and half integral. The transmission as a function of wave length is represented by curve a, Figure 2.

The wave length interval between successive bright bands is inversely proportional to the thickness of the crystal. It is given approximately by setting $\Delta n = 1$ in the equation:

$$\frac{\Delta\lambda}{\lambda} = \frac{\Delta n}{n} \, \frac{1}{\frac{\lambda}{\mu}\frac{\partial\mu}{\partial\lambda} - 1} \qquad (2.3)$$

We now add a second crystal, b_2, and a polarizer p_2 oriented parallel to b_1 and p_1. If $d_2 = 2d_1$, the transmission of the b_2, p_2 combination, represented by curve b, Figure 2, is:

$$\tau_2 = \cos^2 \pi n_2 = \cos^2 \pi 2n_1 \qquad (2.4)$$

The transmission of the whole assembly, b_1, p_1, b_2, p_2, shown in curve c, Figure 2, is therefore:

$$\tau_{12} = \cos^2 \pi n_1 \; \cos^2 \pi 2n_1 \qquad (2.5)$$

A third crystal element b_3, with $d_3 = 2d_2$, followed by the polarizer, p_3, has individual transmission shown in curve d. The transmission of the assembly, b_1 to p_3, is then represented by curve e, Figure 2.

It is evident that further crystal elements and polarizers can be added. The result is the basic type of birefringent filter which will be termed the simple filter. It is comprised of a series of units, each consisting of a plane parallel birefringent element (b-element) followed by a polarizer. All b-elements have surfaces normal to their β-axes and are mounted with their α-axes parallel. All polarizers have their axes parallel to the vibration plane of the entering polarized light at 45° to the α-axes. The thickness of the rth b-element is such that

$$n_r = 2^{r-1} n_1. \qquad (2.6)$$

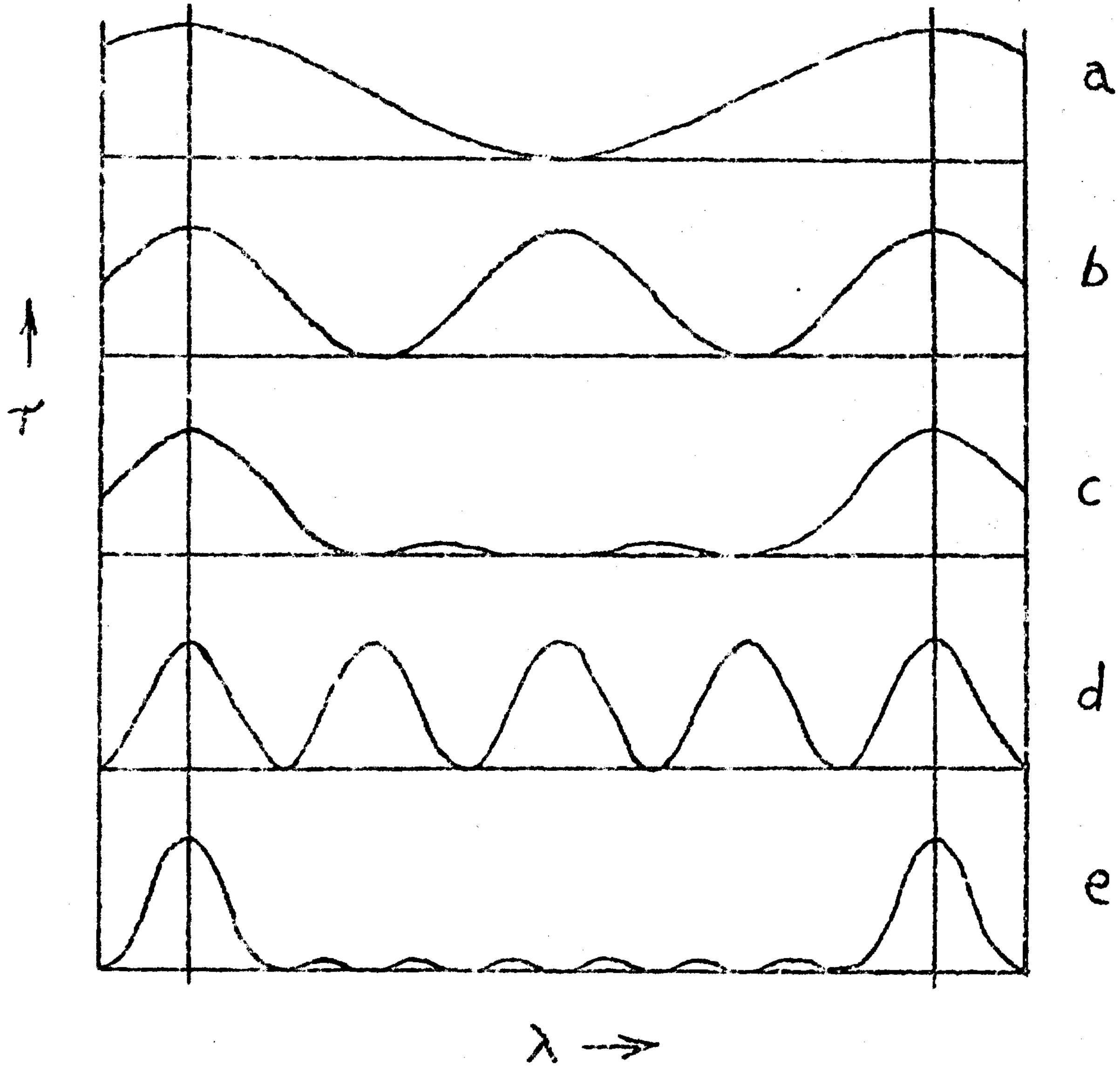


Figure 2. Transmission curves for elements of Figure 1.

a) $b_1\ p_1$; b) $b_2\ p_2$; c) $b_1\ p_1\ b_2\ p_2$;

d) $b_3\ p_3$; e) $b_1\ p_1\ b_2\ p_2\ b_3\ p_3$

The spectrum of light transmitted by the filter consists of a series of widely spaced narrow bands. Their separation is equal to the separation of the transmission maxima of the thinnest element alone, while their effective width is the half width of the maxima of the thickest element alone. For polarized entering light, the transmission of a filter of ℓ b-elements (neglecting absorption in the material of the filter) is:

$$\tau = \cos^2 \pi n_1 \quad \cos^2 \pi 2 n_1 \ . \ . \ . \ \cos^2 \pi 2^{\ell - 1} n_1 \quad (2.7)$$

The quantity n_1 must, of course, be an integer at the wave length of the desired transmission band. Its magnitude should be small enough to separate the adjacent bands sufficiently to permit the isolation of the selected band by means of ordinary filters.

It can readily be shown that the total transmission of flux in an equal energy spectrum is $2^{-\ell}$. Regardless of the width and separation of the bands, the total residual flux transmitted between successive principal maxima in a filter with $\ell > 3$ is a substantially constant fraction (about 0.11) of the flux transmitted in a single band.

The filter at the Climax, Colorado station of the High Altitude Observatory of Harvard University and the University of Colorado has been in satisfactory operation in the observation of solar prominences since early 1943. It is a simple filter of six quartz elements with $n_1 = 23$, $n_6 = 736$, $d_1 = 1.677$ mm and $d_6 = 53.658$ mm and has a transmission band of effective width 4.1 angstroms centered on the H α line of hydrogen (λ6563) at an operating temperature of 35.5° C. Its purpose is to eliminate the overpowering scattered light (continuous spectrum) near the limb of the sun while still transmitting the H α emission from the prominences, which are otherwise completely invisible.

In practice, a filter should either be cemented or immersed in oil to avoid multiple reflections. Initial polarization is usually obtained by mounting a polarizer in front of the first b-element with its axis parallel to the axes of the other polarizers.

In any birefringent crystal both the geometrical dimensions and μ are functions of temperature. The result is a small shift in the wave lengths of the transmission maxima when the temperature changes. In quartz, for instance, $\frac{\Delta \lambda}{\Delta T} = -0.66$ angstrom per degree centigrade in the red.

Hence the temperature of the filter must be controlled with sufficient accuracy to keep the maximum excursions of wave length within tolerable limits. A total range of two tenths of the effective band width is small enough for most purposes.

C. Off Axis Effects in Simple Filters

It is evident that when light traverses a simple filter at an angle to the instrumental axis, the light path through the birefringent material and the velocity difference of the fast and slow waves are altered. The effect is simply to alter the value of n_1 in equation (2.7).

Lyot has calculated the off axis effect for light incident in the two principal planes normal to the α and γ-axes in a biaxial crystal cut with its surfaces normal to the β-axis. Although the equations are not exact, since terms of the fourth and higher degrees in ϕ (the angle of incidence) are neglected, the approximation is excellent for the moderate angles of incidence encountered in the use of filters.

Lyot's equations can be very simply generalized to give the off axis effects for light incident in any plane normal to the surface of the crystal (and parallel, therefore, to the β-axis). Figure 3 represents a block of biaxial crystal with its α, β and γ-axes in the directions indicated. Let polarized light with vibrations in a plane at 45° to the α-axis enter the crystal in the direction (ϕ, θ). Here ϕ is the angle of incidence and θ is the azimuth of the incident plane measured from the α-axis. The light emerges from the crystal in the direction (ϕ, θ) in two polarized components with vibrations very closely parallel to the α and γ-axes. They have a relative retardation, n, which is to be determined as a function of ϕ, θ, and n_0 (where n_0 is the retardation for light entering the crystal from the direction $\phi = 0$).

A consideration of the isochromatic surfaces of biaxial crystals 4* leads to the conclusion that the equations of the curves of constant retardation, n, (written in terms of ϕ and θ) represent hyperbolae if terms in the fourth and higher powers of ϕ are neglected. Their transverse axes are along the α-axis for $\frac{n}{n_0} \gtrless 1$ and along the γ-axis for $\frac{n}{n_0} \lesseqgtr 1$ for crystals in which $\alpha\gamma - \beta^2 \gtrless 1$. The asymptotes are the lines

$$\tan^2 \theta = \frac{\alpha}{\gamma} \tag{3.1}$$

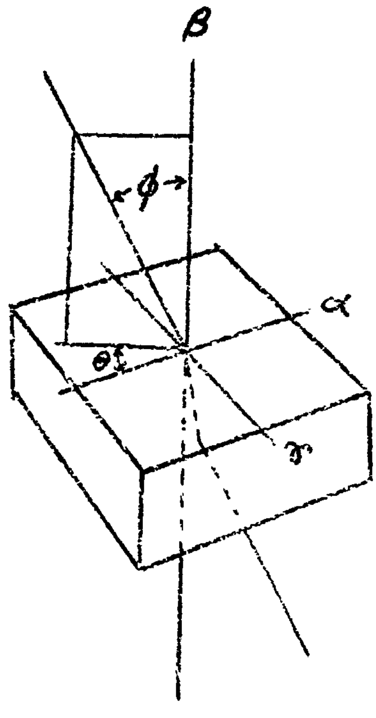


Figure 3. Off axis ray in the crystal coordinate system.

Lyot's equations give the squares of the semi-transverse axes, which are:

$$\phi_0^2 = \left(\frac{n}{n_0} - 1\right) \frac{\gamma}{k} \quad \text{in the plane} \quad \theta = 0$$
$$\phi_{\pi/2}^2 = \left(\frac{n}{n_0} - 1\right) \frac{\alpha}{k} \quad \text{in the plane} \quad \theta = \frac{\pi}{2} \tag{3.2}$$

where

$$k = \frac{\alpha\gamma - \beta^2}{2(\gamma-\alpha)\beta^2} \tag{3.3}$$

We have, therefore, sufficient information to determine both sets of hyperbolas, which can be represented by a single equation:

$$n = n_0 \left[1 + \phi^2 k \left(\frac{\cos^2\theta}{\gamma} - \frac{\sin^2\theta}{\alpha}\right)\right] \tag{3.4}$$

The exact expression for n in uniaxial crystals is readily derived by a straightforward application of Huygens' principle and analytic geometry. Consider a plane parallel uniaxial crystal in a rectangular x, y, z coordinate system with the origin in the first surface. Let it be oriented with its surfaces normal to the z-axis. Let the x-axis be parallel to the crystal optic axis (i.e., parallel to the α-axis in negative crystals or to the γ-axis in positive crystals). Choose units of time and distance to make the velocity of light in space unity. The equation of an entering plane light wave is then:

$$ax + by + cz - t = 0 \tag{3.5}$$

where a, b, and c are the direction cosines of the normal to the wave front and t is the time.

As the wave passes the origin in entering the crystal it initiates a secondary wavelet which expands into an ellipsoid with the equation:

$$\zeta^2 x^2 + \eta^2 y^2 + \nu^2 z^2 - t^2 = 0 \tag{3.6}$$

where ζ, η and ν are reciprocals of the velocities along the x, y and z directions, respectively.

At a given instant, that portion of the plane wave which is inside the crystal coincides with a plane tangent to the ellipsoid of equation (3.6) and containing the line of intersection of the plane wave of equation (3.5) with the first surface of the crystal (i.e., the plane z = 0). The tangent plane through the point (x_1, y_1, z_1) on the ellipsoid is

$$x_1 \zeta^2 x + y_1 \eta^2 y + z_1 \nu^2 z - t^2 = 0. \tag{3.7}$$

The lines of intersection of the planes of equations (3.5) and (3.7) with the first surface of the crystal are respectively:

$$ax + by - t = 0, \qquad z = 0 \tag{3.8}$$

and

$$x_1 \zeta^2 x + y_1 \eta^2 y - t^2 = 0, \qquad z = 0 \tag{3.9}$$

These two lines must coincide, Hence:

$$x_1 = \frac{a}{\zeta^2} t, \qquad y_1 = \frac{b}{\eta^2} t \tag{3.10}$$

Since z_1 must be a point on the ellipsoid of equation (3.6), we find for z_1:

$$z_1 = \frac{t}{v} \sqrt{1 - \frac{a^2}{\zeta^2} - \frac{b^2}{\eta^2}} \tag{3.11}$$

Equations (3.10) and (3.11) define the path of a ray through the origin.

Let d be the thickness of the crystal in the z direction. The time, t_1, when a ray through the origin reaches the second surface is, then:

$$t_1 = \frac{d v}{\sqrt{1 - \frac{a^2}{\zeta^2} - \frac{b^2}{\eta^2}}} \tag{3.12}$$

On emerging from the crystal the plane wave is parallel to the entering wave, with the equation:

$$ax + by + cz - (t - \Delta) = 0 \tag{3.13}$$

At time, t_1, this plane must contain the point (x_1, y_1, d). Hence

$$\Delta = t_1 - (ax_1 + by_1 + cd)$$

The distance, p, of the plane wave of equation (3.13) from the origin is therefore:

$$p = t - \Delta = t - t_1 + ax_1 + by_1 + cd \tag{3.14}$$

or, from equations (3.10) and (3.12):

$$p = t - d\left[\nu\sqrt{1 - \frac{a^2}{\zeta^2} - \frac{b^2}{\eta^2}} - c\right] \qquad (3.15)$$

Now, for the extraordinary wave

$$\zeta = \omega, \quad \eta = \nu = \epsilon$$

and for the ordinary wave

$$\zeta = \eta = \nu = \omega$$

Hence, the distances of the extraordinary and ordinary waves from the origin after their traversal of the crystal can be written, respectively:

$$p_\epsilon = t - d\left[\epsilon\sqrt{1 - \frac{a^2}{\omega^2} - \frac{b^2}{\epsilon^2}} - c\right]$$

$$p_\omega = t - d\left[\omega\sqrt{1 - \frac{a^2+b^2}{\omega^2}} - c\right] \qquad (3.16)$$

The retardation, n, is simply $\frac{P_\omega - P_\epsilon}{\lambda}$, or:

$$n = \frac{n_o}{\epsilon - \omega}\left[\epsilon\sqrt{1 - \frac{a^2}{\omega^2} - \frac{b^2}{\epsilon^2}} - \omega\sqrt{1 - \frac{a^2+b^2}{\omega^2}}\right] \qquad (3.17)$$

Equation (3.17) is the exact expression for the off axis effect in uniaxial crystals. It is readily reduced to the more convenient approximation of equation (3.4). Expanding the radicals, and neglecting fourth and higher powers of a and b, we find:

$$n = \frac{n_o}{\epsilon - \omega}\left[\epsilon - \omega - \frac{\epsilon a^2}{2\omega^2} - \frac{1}{2\epsilon} b^2 \text{ [illegible] } (a^2 + b^2)\right] \qquad (3.18)$$

The direction cosines can be expressed in terms of ϕ and Θ by the transformation:

$$a = \sin\phi \sin\Theta' \qquad b = \sin\phi \cos\Theta'$$

where

$$\theta' = \theta \qquad \text{if } \epsilon - \omega > 0$$

and

$$\theta' = \theta + \frac{\pi}{2} \qquad \text{if } \epsilon - \omega < 0$$

Equation (3.18) becomes, then:

$$n = n_0 \left[1 + \frac{\phi^2}{2\omega}\left(\frac{\cos^2\theta'}{\epsilon} - \frac{\sin^2\theta'}{\omega}\right)\right] \tag{3.19}$$

Equation (3.19) is identical with equation (3.4) for uniaxial crystals.

The corresponding exact equation for biaxial crystals can be derived by the same methods, but the resulting expressions become so lengthy and complicated that it has not seemed worth while to push them through. The accuracy of equation (3.4) is adequate for all practical purposes whether uniaxial or biaxial crystals are considered.

It should be noted that the use of the equations of isochromatic surfaces in the derivation of off axis effects does not lead to an exact result, since they are derived on the inexact assumption that the two components of light polarized at right angles traverse the crystal along identical paths.

The usable field of a given filter is determined by the maximum permissible value of $|n-n_0|$ for the thickest b-element.

The maximum permissible angle of incidence in the Climax filter in the $\theta = \frac{\pi}{2}$ plane is

$$\phi = 0.025 \quad \text{radian}$$

if we require that over the field

$$|n - n_0| \leq 0.1$$

for the thickest b-element.

D. Lyot's Wide Field Filters

The maximum total flux from a given light source that can be squeezed through a filter is roughly proportional to the square of the product of the filter aperture and the maximum usable value of ϕ . The aperture is limited by the sizes of available birefringent crystals, and it is therefore important to find means for obtaining large fields. The most obvious device is to find a birefringent material for which k is very small. Although the author knows of no such material which is available in useful sizes of optical quality, this is a definite possibility which should be investigated further.

Lyot[2*] has described three wide field filters with compound elements made of available materials. They will be referred to as Lyot's first type, second type, and third type filters.

The first type filter differs from the simple filter in having each b-element divided into two equal halves by a cut perpendicular to the β-axis. The second half of each element is rotated about the β-axis until the α-axes of the two components are crossed. A half wave plate is inserted between the components with its α-axis at 45° to the α-axes of the two. It serves to rotate the planes of polarization 90°. Light which enters the first component from the direction (ϕ, θ) enters the second component from the direction $(\phi, \theta + \frac{\pi}{2})$. The retardation introduced by the assembled element is then:

$$n = \frac{1}{2}\left[n\left(\phi,\theta\right) + n\left(\phi,\theta + \frac{\pi}{2}\right)\right]$$

$$= \frac{1}{2}\, n_o \left[1 + \phi^2 k\left(\frac{\cos^2\theta}{\gamma} - \frac{\sin^2\theta}{\alpha}\right)\right]$$

$$+ \frac{1}{2}\, n_o \left[1 + \phi^2 k\left(\frac{\sin^2\theta}{\gamma} - \frac{\cos^2\theta}{\alpha}\right)\right]$$

or

$$n = n_o\left[1 + \phi^2 \frac{k}{2}\left(\frac{1}{\gamma} - \frac{1}{\alpha}\right)\right] \qquad (4.1)$$

The loci of constant retardation are circles with radii larger than the axes of the hyperbolas of a simple filter (in the $\theta = \frac{\pi}{2}$ plane) by a factor of $\sqrt{\frac{2\gamma}{\gamma - \alpha}}$ For a given tolerable value of $|n - n_o|$, the radius of the useful field can be further increased by a factor of $\sqrt{2}$ by setting the retardation at the center of the field at one extreme of the range.

Lyot's first type filter, unlike the simple filter, can be used only over a small range of wave lengths. If the wave length differs greatly from the optimum for which the half wave plates are made, the residual light between transmission bands increases at the expense of light in the bands. The added residual light appears superposed on the field in the form of faint hyperbolic fringes very similar to the fringes produced by the equivalent simple filter. The fringes are lines of constant retardation, n', given by

$$n' = \frac{n_2}{2} k \phi^2 \left(\frac{1}{\gamma} + \frac{1}{\alpha} \right) \cos 2\theta \qquad (4.2)$$

If, however, the filter is either used for one wave length only, or supplied with interchangeable half wave plates for the different spectral regions, its performance is highly satisfactory. This is one of the many instances where the development of an achromatic half wave plate would be very useful.

Lyot's second type wide field filter has compound b-elements of two components of different materials. The quantity k is of opposite sign in the two components, which are mounted with their c-axes parallel. No half wave plates are required.

Let n_1 and n_2 be the retardations due to the first and second components for light entering from the direction $\phi = 0$. The retardation for the assembled element is then:

$$n = n_1 \left[1 + \phi^2 k_1 \left(\frac{\cos^2 \theta}{\gamma_1} - \frac{\sin^2 \theta}{\alpha_1} \right) \right] + n_2 \left[1 + \phi^2 k_2 \left(\frac{\cos^2 \theta}{\gamma_2} - \frac{\sin^2 \theta}{\alpha_2} \right) \right] .$$

$$n = n_0 + \phi^2 \left[\cos^2 \theta \left(\frac{n_1 k_1}{\gamma_1} + \frac{n_2 k_2}{\gamma_2} \right) - \sin^2 \theta \left(\frac{n_1 k_1}{\alpha_1} + \frac{n_2 k_2}{\alpha_2} \right) \right] \qquad (4.3)$$

where now

$$n_0 = n_1 + n_2 . \qquad (4.4)$$

It is evident that while the coefficient of ϕ^2 cannot be made to vanish by any choice of n_1 and n_2, we can obtain circular fringes by eliminating θ . The condition for this is:

$$\frac{n_1}{n_2} = - \frac{k_2}{k_1} \frac{\frac{1}{\gamma_2} + \frac{1}{\alpha_2}}{\frac{1}{\gamma_1} + \frac{1}{\alpha_1}} \qquad (4.5)$$

Equations (4.4) and (4.5) give n_1 and n_2. The retardation of the assembled element can now be written:

$$n = n_o + \phi^2 \left(\frac{n_1 k_1}{\gamma_1} + \frac{n_2 k_2}{\gamma_2} \right) \qquad (4.6)$$

The second type filter can be used over a wide range of wave lengths, although the fringes do not remain strictly circular throughout the range.

Lyot's third type filter generally has the largest field. Each b-element consists of three birefringent components. Two of the components are of the same material and are mounted with their α-axes crossed. The third is of a different birefringent material with a k value opposite in sign to the k value for the first two components. It is mounted with its α-axis parallel to that of one of the first two. By a proper choice of thicknesses it is always possible to make n constant over the whole field within the accuracy of equation (3.4).

Let α_1, β_1, γ_1 and α_2, β_2, γ_2 be the refractive indices of the crystals composing the single component and the two crossed components, respectively. The crystals must be selected to satisfy the condition

$$\alpha_1 \gamma_2 > \gamma_1 \alpha_2$$

Let n_a be the retardation of the single component, and n_b and n_c the retardations of the two components of the same material. Let the α-axes of the a and b components be in the $\Theta = 0$ plane, and the α-axis of the c component in the $\Theta = \frac{\pi}{2}$ plane. Then:

$$\begin{aligned} n = \; & n_a \left[1 + \phi^2 k_1 \left(\frac{\cos^2\Theta}{\gamma_1} - \frac{\sin^2\Theta}{\alpha_1} \right) \right] \\ & + n_b \left[1 + \phi^2 k_2 \left(\frac{\cos^2\Theta}{\gamma_2} - \frac{\sin^2\Theta}{\alpha_2} \right) \right] \qquad (4.7) \\ & - n_c \left[1 + \phi^2 k_2 \left(\frac{\sin^2\Theta}{\gamma_2} - \frac{\cos^2\Theta}{\alpha_2} \right) \right] \end{aligned}$$

If we set $n_a + n_b - n_c = n_o$, and require that the coefficient of ϕ^2 vanish, we find:

$$n_a = \frac{n_o}{A} k_2^2 \left(\frac{1}{\alpha_2^2} - \frac{1}{\gamma_2^2}\right)$$

$$n_b = \frac{n_o}{A} k_1 k_2 \left(\frac{1}{\gamma_1\gamma_2} - \frac{1}{\alpha_1\alpha_2}\right) \tag{4.8}$$

$$n_c = \frac{n_o}{A} k_1 k_2 \left(\frac{1}{\alpha_1\gamma_2} - \frac{1}{\gamma_1\alpha_2}\right)$$

where

$$A = \begin{vmatrix} \frac{k_1}{\gamma_1} & \frac{k_2}{\alpha_2} & \frac{k_2}{\gamma_2} \\ \frac{k_1}{\alpha_1} & \frac{k_2}{\gamma_2} & \frac{k_2}{\alpha_2} \\ 1 & -1 & 1 \end{vmatrix} \tag{4.9}$$

The retardation of the assembled element for any direction (ϕ, Θ) is then:

$$n = n_a + n_b - n_c = n_o \tag{4.10}$$

The third type filter like the second type, can be used over a wide range of wave lengths. The coefficient of ϕ^2, however, will generally vanish accurately at only one wave length.

In designing a wide angle filter it is not usually necessary to make the thinner elements compound. Their transmission bands are so broad that the slight shift in wave length for off axis rays is negligible in comparison. If the higher order compound elements are made of two materials, however, it may not be possible to use transmission bands in widely separated regions of the spectrum, because the dispersions of different materials are generally not strictly proportional. If the r th element is the thickest simple element, $\frac{n_{r+1}}{n_r} = 2$ at only one wave length.

II. RECENT MODIFICATIONS OF THE BIREFRINGENT FILTER

The following sections are devoted to the theory of various modifications of birefringent filters which have been recently developed.

A. The Split Element Filter

The split element filter resembles Lyot's first type filter, and shares its wide field characteristics. The half wave plates, however, are replaced by birefringent elements, and successive polarizers are crossed. After the initial polarization, it requires only half as many polarizers as the equivalent simple filter. The result is a considerable reduction in absorption and scattered light if film polarizers are used, or a notable saving in bulk and expense if polarizing prisms are used.

The split element filter has already been described briefly[5*]. A more detailed account of its theory is given here.

A single unit of the split element filter (which would be mounted between crossed polarizers) is shown schematically in Figure 4. The x, y and z-axes constitute a rectangular coordinate system. The positive r and s-axes in the xy plane bisect the angles between the positive x and y and the positive y and negative x directions, respectively. The unit consists of a split element with components m and q, and a simple element, p, sandwiched between m and q. They are all mounted with β-axes parallel to the z-axis. The γ-axes are aligned parallel to the x, r and y directions, respectively, in the m, p and q components. Let the thicknesses of m, p and q be d_m, d_p and d_q, and let the unit of time be the vibration period of the light.

Assume that the entering light is polarized in the r plane. The transmissions of the unit for emerging light polarized in the r plane and s plane are to be determined.

The vibration of the entering light is:

$$r = a \sin 2\pi t \qquad (5.1)$$

This can be resolved along the x and y directions, giving:

$$x = \frac{a}{\sqrt{2}} \sin 2\pi t$$

$$y = \frac{a}{\sqrt{2}} \sin 2\pi t \qquad (5.2)$$

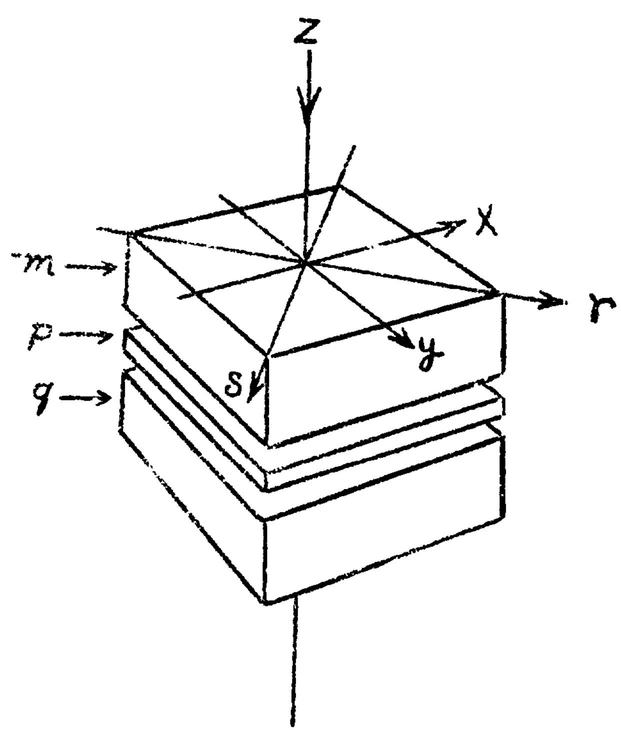


FIGURE 4

Birefringent components of a single unit of a split element filter

In traversing m, a phase difference is introduced and the vibration of the emerging light is

$$x_m = \frac{a}{\sqrt{2}} \sin 2\pi \, (t - d_m \gamma)$$

$$y_m = \frac{a}{\sqrt{2}} \sin 2\pi \, (t - d_m \alpha) \qquad (5.3)$$

The resultant disturbance along the r and s axes is:

$$r_m = a \cos \pi n_m \quad \sin 2\pi t'$$

$$s_m = a \sin \pi n_m \quad \cos 2\pi t' \qquad (5.4)$$

where

$$t' = t - \frac{d_m}{2\lambda} (\alpha + \gamma)$$

In the traversal of p, an additional phase difference is introduced:

$$r_p = a \cos \pi n_m \quad \sin 2\pi \left(t' - \frac{d_p}{\lambda} \gamma\right)$$

$$s_p = a \sin \pi n_m \quad \cos 2\pi \left(t' - \frac{d_p}{\lambda} \alpha\right) \qquad (5.5)$$

Resolving this vibration along the x and y axes and adding the phase difference due to transmission through q, we obtain:

$$x_q = \frac{a}{\sqrt{2}} \cos \pi n_m \quad \sin 2\pi \left(t' - \frac{d_p}{\lambda} \gamma - \frac{d_q}{\lambda} \alpha\right)$$
$$- \frac{a}{\sqrt{2}} \sin \pi n_m \quad \cos 2\pi \left(t' - \frac{d_p}{\lambda} \alpha - \frac{d_q}{\lambda} \alpha\right) \qquad (5.6)$$

$$y_q = \frac{a}{\sqrt{2}} \cos \pi n_m \quad \sin 2\pi \left(t - \frac{d_p}{\lambda} \gamma - \frac{d_q}{\lambda} \gamma\right)$$
$$+ \frac{a}{\sqrt{2}} \sin \pi n_m \quad \cos 2\pi \left(t' - \frac{d_p}{\lambda} \alpha - \frac{d_q}{\lambda} \gamma\right)$$

To determine the final transmission through a polarizer with its axis along either the r or s direction, we must resolve this vibration along the r and s axis:

$$
\begin{aligned}
r_q &= a \cos \pi n_m \cos \pi n_q \sin 2\pi \left(t'' - \frac{d_p}{\lambda} \gamma\right) \\
&\quad + a \sin \pi n_m \sin \pi n_q \sin 2\pi \left(t'' - \frac{d_p}{\lambda} \alpha\right) \\
s_q &= a \sin \pi n_m \cos \pi n_q \cos 2\pi \left(t'' - \frac{d_p}{\lambda} \gamma\right) \\
&\quad - a \cos \pi n_m \sin \pi n_q \cos 2\pi \left(t'' - \frac{d_p}{\lambda} \alpha\right)
\end{aligned} \tag{5.7}
$$

where

$$t'' = t - \frac{d_m + d_q}{2\lambda} \left(\alpha + \gamma\right)$$

Let the emergent amplitudes be A_r and A_s. The transmissions in the r and s vibration planes are, then:

$$
\begin{aligned}
T_r &= \frac{A_r^2}{a^2} = \cos^2 \pi (n_m - n_q) - \sin 2\pi n_m \sin 2\pi n_q \sin^2 \pi n_p \\
T_s &= \frac{A_s^2}{a^2} = \sin^2 \pi (n_m - n_q) + \sin 2\pi n_m \sin 2\pi n_q \sin^2 \pi n_p
\end{aligned} \tag{5.8}
$$

In the split element filter the m and q components are made of equal thicknesses. Hence

$$n_m = n_q \, .$$

If we let

$$n_j = 2\, n_m = 2\, n_q$$

equations (5.8) reduce to:

$$
\begin{aligned}
T_r &= 1 - \sin^2 \pi n_j \sin^2 \pi n_p \\
T_s &= \sin^2 \pi n_j \sin^2 \pi n_p
\end{aligned} \tag{5.9}
$$

The transmission of an element of Lyot's first type filter is T_r in equations (5.9) if we set $n_p = \frac{1}{2}$.

The transmission of a unit of the split element filter is T_s. A split element filter of ℓ elements has exactly the same off axis characteristics as a filter of Lyot's first type with the first $\frac{\ell}{2}$ elements simple, and the $\frac{\ell}{2}$ thicker elements compound. Whether the field is limited to the simple elements or the compound elements depends upon whether or not $\frac{n_\ell}{n_{\ell/a}} \cdot \frac{\gamma-a}{2\gamma}$ is greater or less than 1. If the simple elements limit the field, they can, of course, be made compound in any of Lyot's three types.

The transmission of an assembled split element filter composed of two-element units between crossed polarizers is:

$$\mathcal{T} = \sin^2 \pi n \;\; \sin^2 \pi n_2 \; . \; . \; . \; . \; \sin^2 \pi n_\ell \tag{5.10}$$

Since transmission bands occur only at wave lengths for which all the n's are half integral, the n's cannot be simply proportional to the powers of 2. If we let $n = n' + \frac{1}{2}$ at the wave length of a particular band, the best we can do is to make the values of n' proportional to the powers of 2. Thus

$$n_r = 2^{r-1} n_1' + \frac{1}{2} \tag{5.11}$$

The transmission can then be written

$$\mathcal{T} = \cos^2 \pi n_1' \;\; \cos^2 \pi 2 n_1' \; . \; . \; . \; \cos^2 \pi \, 2^{\ell-1} n_1' \tag{5.12}$$

Unfortunately equation (5.11) can be strictly valid at only one wave length, and the usefulness of the filter is restricted to a limited spectral region in the neighborhood of that wave length. This is a second instance where achromatic half wave plates would be handy. If the r th element of the filter were made to give a retardation $n_r' = 2^{r-1} n_1'$, the addition of an achromatic half wave plate (two quarter wave plates for split elements) would satisfy equation (5.11) at all wave lengths.

The thought will doubtless have occurred to the reader that the middle element in each unit of a split element filter could itself be split, and a third element inserted between the halves. This plan does not work theoretically, and so far no arrangement has been found which allows more than two elements in a unit between successive polarizers.

B. Filters of Adjustable Wave Length

It is obvious that the usefulness of the birefringent filter is enormously enhanced if a transmission maximum can be adjusted to center on any desired wave length. The fine adjustment resulting from the control of temperature is generally quite inadequate as it has a range of only a few angstroms (although Lyot found that with the aid of temperature control it is possible to bring no less than six of the maxima of a quartz filter into coincidence with lines of major importance in the solar spectrum).

The obvious method of controlling the wave length of the transmission bands is by means of elements of variable thickness, made of pairs of wedges which can be adjusted with respect to each other like the components of a Babinet compensator. It is then possible to set

$$n_1 = \text{an integer}$$

and

$$n_r = 2^{r-1} n ,$$

at any chosen wave length. Such an arrangement is perfectly feasible and works equally well at all wave lengths. In the split element filter, both halves of the split element must, of course, be adjustable, since $n_m - n_q = 0$. The range of variation in thickness need be only sufficient to shift the principal transmission maxima of the filter through a range equal to their separation. With a proper choice of wedge angles, all the movable wedges can be mounted and adjusted as a single unit.

Although theoretically excellent, the variable thickness filter requires considerable mechanical refinement, and one wedge in each element must have an aperture much larger than the instrumental aperture (a matter of importance in filters of large aperture). The use of phase shifters for wave length adjustment is simpler and, for most purposes, equally satisfactory. If achromatic phase shifters can be devised, they will give results as theoretically perfect as variable thickness.

Suppose we equip each b-element of a filter with a phase shifter which permits the addition of a small controllable phase difference, $2\pi\zeta$, to the phase difference, $2\pi n$, introduced by the b-element. The transmission of the filter is then

$$\tau = \prod_{r=1}^{r=\ell} \cos^2 \pi \left(n_r + \zeta_r\right) \qquad (6.1)$$

Again, with the split element filter, the added phase difference must be divided equally between the two halves of the split elements to keep $n_m - n_q = 0$. A transmission maximum of the filter can then be centered on any given wave length, λ, by adjusting ζ until $n + \zeta$ is an integer for each element. This is always possible if ζ can be adjusted over the range $-\frac{1}{2}$ to $+\frac{1}{2}$. If the phase shifter is achromatic, i.e., ζ is independent of wave length at a given setting, the result is merely a shift of the transmission curve of the filter along the spectrum and its performance is equally good at all wave length settings. If, on the other hand, δ is a function of wave length, the spacings of the transmission maxima of a given element are altered. Hence the relative positions of the transmission maxima and minima of the different elements depart more and more from exact superpositon as the wave length departs from λ_1. The result is an increase in the residual light transmitted in the intervals between principal maxima of the filter as $|\lambda - \lambda_1|$ increases.

Lyot[3*] and Billings[6*] have both made numerical calculations of the additional residual light resulting from the use of non achromatic phase shifters. They concluded that over a reasonable wave length range (which can readily be isolated with glass or gelatine filters) the increase in residual light is negligible. The adjustment of wave length with phase shifters is therefore a practical possibility whether the phase shifters are achromatic or not.

Several forms of variable phase shifters have been proposed.

Lyot[3*] made elements of variable thickness like those described above for the variable thickness filter, but with the difference that the range of adjustment of retardation was restricted to one wave length.

Billings[6*] made an experimental filter with photo elastic phase shifters composed of sheets of polyvinyl butyrate under adjustable tension.

While both these arrangements give a satisfactory wave length adjustment, they are tedious to use. Ordinarily each element must be individually adjusted. The alternative is a complicated mechanical synchronization of the adjustments of all the elements, which would make operation with a single control feasible. Without some such arrangement it would be impossible to vary the wave length continuously.

A much more promising approach is the use of the electro optical phase shifters discussed by Billings[6*]. A plate of the uniaxial crystal ammonium di hydrogen phosphate ($NH_4H_2PO_4$), known commercially as PN, cut perpendicular to the optic axis and mounted between transparent electrodes, becomes biaxial and exhibits a retardation when a potential difference is applied to the electrodes. The retardation is proportional to the potential difference and is independent of the thickness of the PN plate. A filter made with a Billings plate added to each element (to each half of the split elements in the split element filter) could be adjusted electrically, and the problem of synchronizing the phase shifts of successive elements would be relatively simple. At the present writing Dr. Billings is actively engaged in the development of such electrically tunable filters.

All three tuning methods have one difficulty in common. It is impracticable to push the phase shift beyond a very limited range. If a range from $-\pi$ to $+\pi$ is adopted, a continuous variation of wave length involves a discontinuous adjustment of each phase shifter. The phase shift must progress smoothly from $-\pi$ to $+\pi$ (at a rate proportional to the thickness of the associated b-element) and then jump back to $-\pi$. For most purposes there may be no serious disadvantage in this. If, however, the filter is to be used for spectrophotometric work, for example, it may be very difficult to avoid a spurious bump in the filter transmission every time a phase shifter passes a point of discontinuity, even with the electrical tuning. For such special purposes phase shifters composed of rotating fractional wave plates can be used. They have already been described briefly[5*]. A fuller account of their theory is given here.

The specific problem is to devise a combination of fractional wave plates which will alter the phase difference between the vibrations along two mutually perpendicular axes, x and y, by any chosen amount, without altering their amplitudes. At a given wave length this is equivalent to a variable thickness of birefringent material with its γ-axis along the x or y direction. Such an arrangement is shown at a, Figure 5. It consists of two quarter wave plates. The first is fixed with its γ-axis along the r axis (at 45° to the x-axis). The second can be rotated around the instrumental axis. At a given setting its γ-axis lies along the r' direction at angle ρ to the r direction.

The vibration of the light entering the system is generally represented by

$$x = b \sin 2\pi t$$

$$y = c \sin 2\pi (t + \sigma) \qquad (6.2)$$

Resolving this vibration along the r and s axes and adding a phase difference of $\frac{\pi}{2}$ introduced by the first quarter wave plate, we obtain for the emerging vibration:

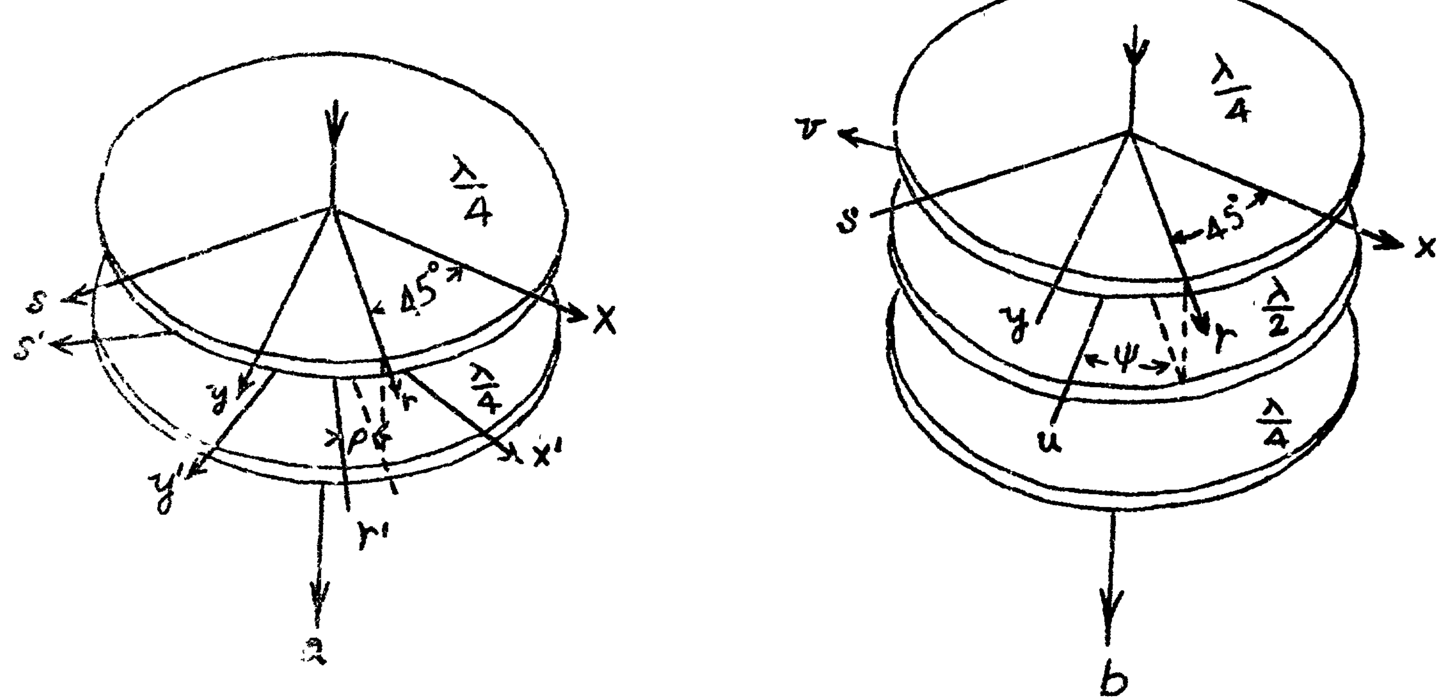


FIGURE 5

a) Fiuse shifter of two quarter wave plates.

b) Phase shifter of one half wave and two quarter wave plates.

$$r = \frac{b}{\sqrt{2}} \sin 2\pi t + \frac{c}{\sqrt{2}} \sin 2\pi(t+\sigma)$$

$$s = -\frac{b}{\sqrt{2}} \cos 2\pi t + \frac{c}{\sqrt{2}} \cos 2\pi(t+\sigma) \qquad (6.3)$$

Resolving this vibration along the r' and s' axes and adding another phase difference of $\frac{\pi}{2}$ introduced by the second quarter wave plate, we obtain:

$$r' = \frac{b}{\sqrt{2}} \sin \left[2\pi t - \rho\right] + \frac{c}{\sqrt{2}} \sin \left[2\pi(t+\sigma) + \rho\right]$$

$$s' = \frac{b}{\sqrt{2}} \sin \left[2\pi t - \rho\right] - \frac{c}{\sqrt{2}} \sin \left[2\pi(t+\sigma) + \rho\right] \qquad (6.4)$$

Finally, if we resolve this vibration along the x' and y' axes, at an angle of $\rho + \frac{\pi}{4}$ to the x and y axes, we obtain for the emerging vibration:

$$x' = b \sin \left[2\pi t - \rho\right]$$

$$y' = c \sin \left[2\pi(t+\sigma) + \rho + \pi\right] \qquad (6.5)$$

A comparison of equations (6.2) with (6.5) shows that while the emerging amplitudes along x' and y' are the same as the entering amplitudes along x and y, the phase difference has been increased from $2\pi\sigma$ to $2\pi\sigma + 2\rho + \pi$, i.e., the phase shift, $2\pi\xi$, is

$$2\pi\xi = \pi + 2\rho \qquad (6.6)$$

Obviously the phase difference can be set to any desired value by adjusting ρ.

This two element phase shifter has the disadvantage that the x' and y' axes rotate with the second quarter wave plate. For some applications this is no inconvenience but in others it renders this phase shifter useless. The x' and y' axes can be restored to parallelism with the x and y axes by the addition of a rotatable half wave plate, which has the property of reflecting any polarization figure in its γ-axis.

The most convenient system, shown at b, Figure 5, consists of two fixed quarter wave plates with the rotatable half wave plate sandwiched between them. Suppose the γ-axes of both quarter wave plates are in the r direction, while the γ-axis of the half wave plate is along the u direction at an angle ψ to the r direction.

The vibration emerging from the first quarter wave plate is given by equation (6.3). Resolving this vibration along the u and v-axes, and adding a phase difference of π, we obtain for the vibration emerging from the half wave plate:

$$u = \frac{b}{\sqrt{2}} \sin\left[2\pi t - \psi\right] + \frac{c}{\sqrt{2}} \sin\left[2\pi (t + \sigma) + \psi\right] \qquad (6.7)$$

$$v = \frac{b}{\sqrt{2}} \cos\left[2\pi t - \psi\right] - \frac{c}{\sqrt{2}} \cos\left[2\pi (t + \sigma) + \psi\right]$$

Resolving this vibration again along the r and s axes, and adding a phase difference of $\frac{\pi}{2}$, we obtain for the vibration emerging from the second quarter wave plate:

$$r = \frac{b}{\sqrt{2}} \sin\left[2\pi t - 2\psi\right] + \frac{c}{\sqrt{2}} \sin\left[2\pi (t + \sigma) + 2\psi\right] \qquad (6.8)$$

$$s = \frac{b}{\sqrt{2}} \sin\left[2\pi t - 2\psi\right] + \frac{c}{\sqrt{2}} \sin\left[2\pi (t + \sigma) + 2\psi\right]$$

Finally, resolving this vibration along the original x and y axes, we find

$$x = b \sin\left[2\pi t - 2\psi\right]$$

$$y = c \sin\left[2\pi (t + \sigma) + 2\psi\right] \qquad (6.9)$$

The phase shift introduced by the three element system is, therefore

$$2\pi\xi = 4\psi \qquad (6.10)$$

The principal advantage in the use of fractional wave plate phase shifters in birefringent filters is in the possibility of a continuous variation of wave length without discontinuities in the adjustment of the moving elements. Since ρ or ψ can be increased or decreased indefinitely, $2\pi\xi$ is not restricted as it is in the other types of phase shifters discussed above.

It should be noted that the fractional wave plate phase shifter is in a sense achromatic, since ξ is independent of the wave length for a given value of ρ or ψ — a very desirable property (see the discussion following equation 6.1). With ordinary quarter and half wave plates, however, this advantage is somewhat illusory. Their usefulness is limited to the rather restricted region of the spectrum where their retardations are very close to quarter wave and half wave. This is another application where the desirability of achromatic fractional wave plates is evident.

If continuity of adjustment over a large range of the spectrum is a necessity, the fractional wave plates themselves could be made adjustable. The addition of an electro optical Billings plate to each fractional wave plate would perhaps be the simplest method. A relatively moderate potential applied to the Billings plate would then adjust the retardation accurately to a half wave or quarter wave at the wave length of the transmission band of the filter. This seems a rather desperate measure, however.

The construction of the fractional wave plate phase shifters is considerably simplified when they are used in birefringent filters. Some of the quarter wave plates simply take the form of an addition to the thickness of the birefringent elements. In instances where the γ-axis of a quarter wave plate is parallel or perpendicular to the axis of an immediately following polarizer, it is evident that the polarizer utilizes only one component of the vibration emerging from the quarter wave plate. The $\frac{\pi}{2}$ phase difference therefore serves no real purpose, and the quarter wave plate can be omitted.

Consider first an element of a simple filter. Suppose the b-element, oriented with its γ-axis along the x direction, is followed by a quarter wave plate with its γ-axis along the r direction. If we let $b = c\frac{d}{\sqrt{2}}$, $t = t' - \frac{d}{2\lambda}\mu$, and $\sigma = \frac{d}{\lambda}\mu$, equation (6.3) for the vibration emerging from the quarter wave plate reduces to:

$$r = a \cos \pi n \sin 2\pi t' \qquad (6.11)$$

$$s = a \sin \pi n \sin 2\pi t'$$

This is a linear vibration at an angle of πn to the r-axis. We can omit the second quarter wave plate and let the light enter a polarizer with its plane of polarization at angle ρ to the r axis. The transmission of the assembly is then

$$\Upsilon = \cos^2 (\pi n - \rho) \qquad (6.12)$$

By adjusting ρ (i.e., by rotating the polarizer) until $n_\lambda - \frac{\rho}{\pi}$ = an integer, we can set $\Upsilon = 1$ for any chosen wave length.

Lyot[39] has utilized this device to effect a slight shift in the wave length of the transmission band of his filter.

He used a quarter wave plate with the last (thickest) element, and provided for the rotation of the final polarizer. The same method can be applied to the whole filter, however.

An adjustable simple birefringent filter would consist, then, of a series of units shown at a, Figure 6, each composed of a polarizer, a birefringent element with its γ-axis at 45° to the axis of the polarizer, and a quarter wave plate with its γ-axis parallel to the axis of the polarizer. The three parts of each unit remain fixed with respect to each other, but the unit itself must be rotatable around the instrumental axis. The angle ρ_r is then the angle between the γ-axis of the r th quarter wave plate and the axis of the immediately following polarizer. The birefringent elements have the same thickness as in the non-adjustable filter. The transmission of the whole is:

$$\tau = \cos^2(\pi n_1 - \rho_1)\cos^2(\pi 2n_1 - \rho_2)\ \cdots\ \cos^2(\pi 2^{l-1} n_1 - \rho_l) \qquad (6.13)$$

and

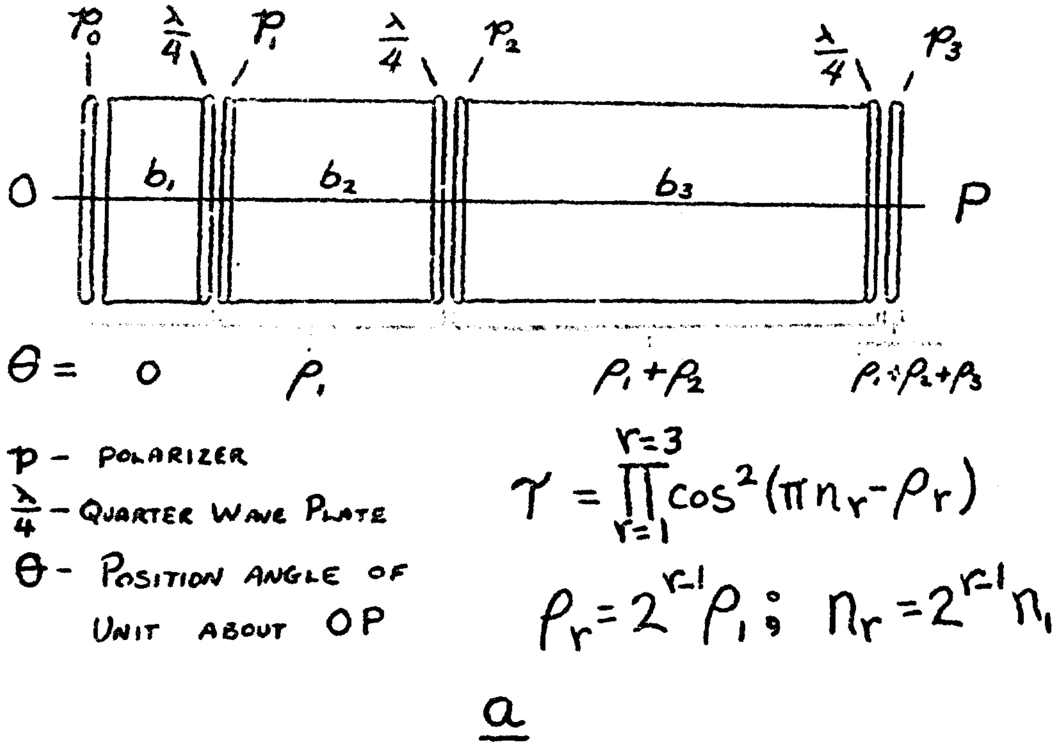
$$\rho_r = 2^{r-1}\rho_1 \qquad (6.14)$$

Since the values of ρ are proportional to the powers of 2, it is a relatively simple matter to devise a gear train by which the wave length of the transmission band can be adjusted with a single control knob. A continuous variation of wave length now involves no discontinuity in the adjustment of the various units, since ρ can be made to increase or decrease indefinitely.

Matters are somewhat more complicated in the split element filter. The wide field characteristics depend upon the m and q components being crossed. Hence the phase shifts must be accomplished without any relative rotation of the two. Various arrangements are possible, some of which involve rotation of the center p-elements, or rotation of the unit as a whole with respect to the polarizers, or both. However, the unit shown at b, Figure 6, is as simple as any.

The orientation of each element is indicated in the diagram by the short line above it for the fixed elements, or by the symbol ψ_{mq} or ψ_p for the adjustable half wave plates. The angle ψ_{mq} or ψ_p is the angle between the γ-axis of the half wave plate and the γ-axis of the preceding quarter wave plate. The second quarter wave plates following m and p are indicated as an addition to the thicknesses of p and q, while that following q has been omitted, since its γ-axis would be parallel to the axis of the following polarizer. The transmission of a split element filter composed of such units is

$$\tau = \prod_{r=1}^{r=l} \cos^2\left(\pi n_r - 2\psi_r - \frac{\pi}{2}\right) \qquad (6.15)$$



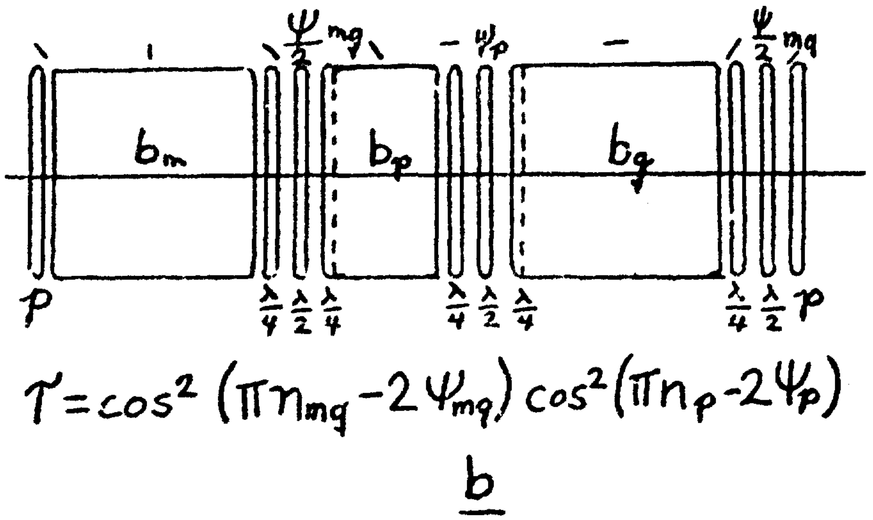


FIGURE 6

a) Simple filter of three elements with quarter wave plate phase shifters.

b) One unit of a split element filter with fractional wave plate phase shifters.

It should be noted here that the built-in quarter wave plates which are added to the thicknesses of the p and q-elements are not included in the calculation of n for these elements.

The values of ψ_k should be proportional to n_r in equation (6.15). Hence if $n_r = 2^{r-1}(n_1 - \frac{1}{2}) + \frac{1}{2}$ as in the non-adjustable split element filter, the ψ's are proportional to large odd numbers, and the problem of synchronizing the rotations of the half wave plates becomes complicated (but not at all impossible). If, on the other hand, the n's are made proportional to the powers of two, the phase changers can compensate for the subtraction of $\frac{1}{2}$ from each value of n in addition to their normal function. Then

$$n_r = 2^{r-1} n_1 \tag{6.16}$$

and

$$2\psi_r = \frac{\pi}{2} + 2^{r-1} (2\psi_1 - \frac{\pi}{2}) \tag{6.17}$$

Since a rotation of the zero point from which angle ψ is measured to $\frac{\pi}{4}$ reduces this equation to

$$2\psi'_r = 2^{r-1} (2\psi'_1) \tag{6.18}$$

it is evident that the variable parts of the ψ's are proportional to the powers of two, and the problem of synchronization becomes relatively simple.

The synchronization of the other types of phase shifters (variable thickness, photo elastic, or electro optical) is similarly simplified in a split element filter by constructing it with n's proportional to powers of two. Equations (6.14) and (6.17) apply if we substitute $\pi\xi$ for 2ψ.

A final remark about filters of adjustable wave length seems worth while. The birefringent elements need not be made to any exact thicknesses as in the fixed wave length filters. It is desirable, but not necessary, to preserve the relation $n_r = 2^{r-1}n_1$ as closely as possible since the synchronization of the various adjustments is then easier. There is no necessity, however, for n_1 to be an integer for any specified wave length. This simplifies the construction somewhat. If $\mu \leq 0.03$ the thicknesses of the elements can be adjusted with sufficient accuracy by mechanical measurements alone. The error tolerance in thickness is inversely proportional to μ and is about ± 0.001 mm for $\mu = 0.03$.

III Materials and Applications

A. Materials for Birefringent Filters

For the benefit of potential builders of birefringent filters a brief discussion of available materials is given below. It must be emphasized that the list given is certainly far from complete. The author simply lists materials which have come to his attention and either have been successfully used, or look promising. Unfortunately, lack of time has prevented a really thorough search for suitable and available materials, and it would be surprising if some very useful ones had not been overlooked.

Some of the desirable properties of crystals for birefringent filters are a large value of μ with a small temperature coefficient; a high degree of hardness, chemical stability, and insolubility in water; high transparency in the region of the spectrum for which the filter is to be used; and availability in large pieces of high optical quality.

For filters with band widths of 3 angstroms or more quartz is an ideal material. It is excellent on all counts except for its rather small value of μ $(=0.009)$. The birefringent elements of all the astronomical filters now in operation are made of quartz except for the final element of Lyot's filter, which is calcite.

Calcite would be excellent for elements of large n-values if it were readily available in large sizes. Unfortunately it is so difficult to obtain that its general use in filters is probably impossible. While it is not as easily ground and polished as quartz, it presents no real difficulty. $\mu = 0.17$.

Gypsum occurs naturally in large crystals and should be readily available. Its birefringence is similar to that of quartz, and it should be useful in the same places. Unfortunately, it is quite soft and might be difficult to polish. $\mu = 0.009$.

Ammonium di-hydrogen phosphate has excellent optical characteristics, although it is sensitive to pressure and must be mounted with care. It is available in large sizes. Its optical working has proved rather difficult, though not impossible, and its high solubility in water necessitates careful protection from atmospheric moisture. $\mu = 0.045$.

Ethylene diamine tartrate has promising optical characteristics accompanied by the disadvantages of high solubility in water and softness. The author knows of no attempts to polish it, but it would probably be quite difficult. It is available in large sizes. $\mu = 0.084$

Sodium nitrate has a larger μ value than calcite, and should be useful for elements of large n-values. However, it is very soluble in water and difficult to work. At present it is not available in large sizes with the necessary homogeneity. $\mu = 0.25$.

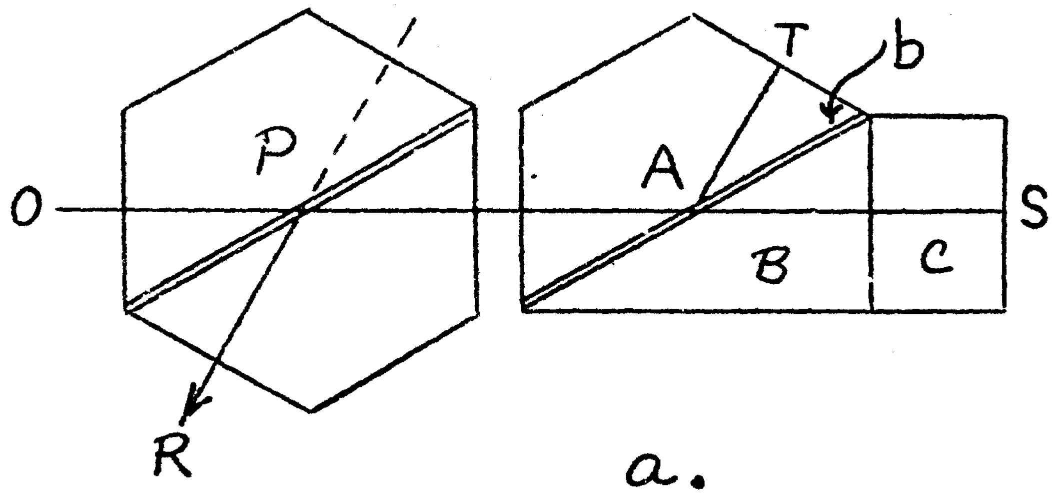
S. Polarizing interferometer filters

An account of birefringent filters should not be closed without some mention of the polarizing interferometer, a device which has the effect of an impossibly thick birefringent element. It offers the possibility of filters of very high resolution with band widths in the range of hundredths or thousandths of an angstrom. The advantages of the polarizing over the usual forms of interferometers is in the possibility of an accurate and stable control of the wave lengths of transmission maxima (by means of phase shifters), and a high light efficiency.

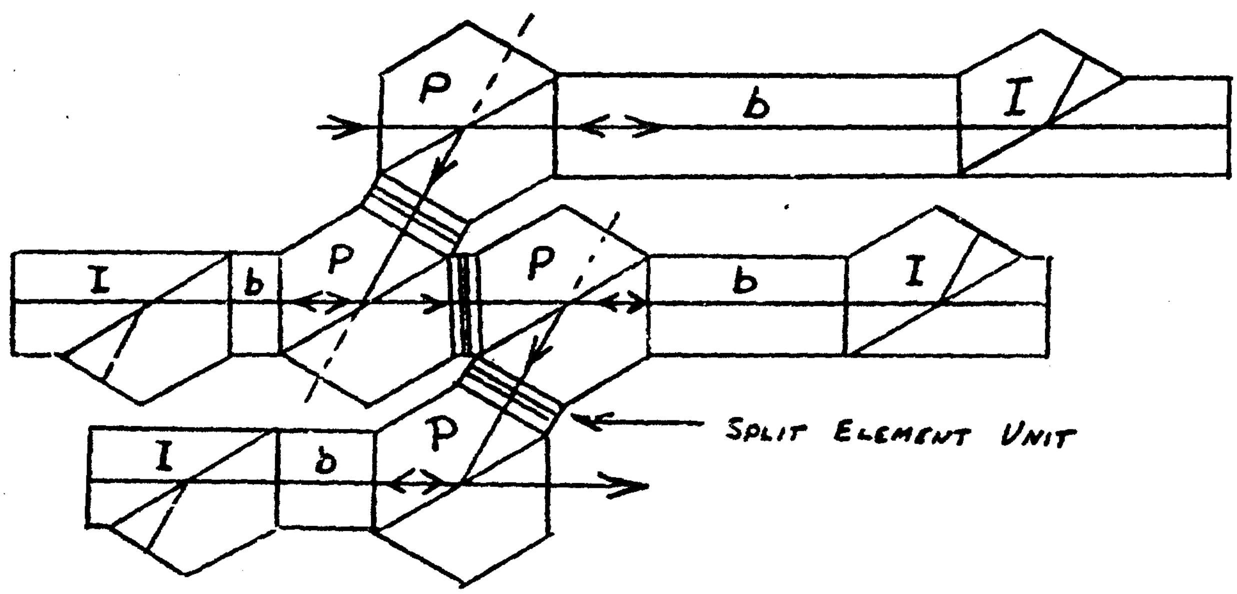
The essential feature of the polarizing interferometer is that the emerging light consists of two coherent sets of waves which differ in phase (due to a path difference) and are polarized at right angles to each other. The effect is similar to that of a birefringent element, and a series of polarizing interferometers can be used exactly like a series of birefringent elements to construct a filter. The wave length of the transmission band can be controlled with adjustable phase shifters, and interferometers can be sandwiched between birefringent elements to form split element units.

The advantage of the polarizing interferometer over a simple birefringent element is that very large values of n can be obtained in a comparatively compact element. The saving in bulk may not be important, but the difficulty of obtaining birefringent material in very great thicknesses is. An element of calcite, for instance, must be about eleven times as thick as a path difference in glass. The principal disadvantage is the expense of construction common to all interferometers of the split amplitude class. The field is small for large values of n, and while it is theoretically quite simple to make a birefringent field compensator, it is impractical because the thickness of birefringent material required nullifies the advantage of compactness.

Many forms of polarizing interferometers are possible. One type which is well adapted for the construction of filters is shown at a, Figure 7. It is a modified solid Michelson interferometer with a polarizing beam splitter. It consists of two glass prisms, A and F, with a very thin slip, b, of sodium nitrate (or other highly birefringent material) cemented between them with its optic axis normal to the surface. If the angles are properly



a.



b.

FIGURE 7

a) One form of polarizing interferometer.

b) High resolution filter composed of polarizing interferometers and birefringent elements.

chosen, the b-layer totally reflects the light vibrating in the plane of the drawing and transmits the light vibrating at right angles to it. A spacer element, C, introduces a path difference. Surfaces S and T are silvered or aluminized. Light which enters in the direction OS, emerges in the reverse direction, SO, in two components polarized at right angles, with a phase difference given by:

$$2\pi n = 4\pi \frac{\mu'}{\lambda} d_c \cos\phi \tag{7.1}$$

where μ' is the refractive index and ϕ is the angle of incidence on S and T. The prism P (constructed like A,B) has the double function of polarizing entering light and separating out the desired part of the emerging light. It is shown in an incorrect orientation for simplicity in drawing. Actually prism P is rotated about the OS direction, to bring its axis to an angle of 45° to that of prism AB. The transmission of the whole assembly for light emerging in the R direction is then

$$\tau = \sin^2 \pi n. \tag{7.2}$$

The remainder of the light emerges along SO.

The most serious difficulty in the construction of such an interferometer is the optical working and cementing of the b-layer to the required accuracy. The orientation of the S and T surfaces with respect to each other is not so critical, since a slight misalignment can be compensated by a thin wedge of birefringent material between prism P and the interferometer.

One method of using polarizing interferometers combined with birefringent elements in a filter is shown schematically at b, Figure 7. Between each polarizer, P, and the following interferometer, I, is a b-element, which constitutes the m (for entering light) and q (for emerging light) components of a split element. The interferometer then takes the place of the p component. Between successive polarizers are purely birefringent split element units. The assembly includes 4 interferometers, 4 polarizing prisms and 10 b-elements. The interferometers and b-elements should be equipped with phase shifters (not shown). As an example, the interferometers might have retardations of 245,760; 122,880; 61,440; 30,720; and the b-elements, retardations from 15360.5 to 30.5 at $\lambda = 5000$ angstroms. The system would transmit bands of about 0.01 angstrom effective width, spaced about 150 angstroms apart. By adjusting the phase shifters a selected band could be made to scan the spectrum.

If the light transmitted by the filter is received on a photoelectric cell, its output gives a high resolution spectrophotometric curve of the entering light.

Such a filter would be preferable to a grating spectrograph for spectrophotometric purposes, because, in spite of its small field (maximum usable ∅ about 0.0012 radian), it can be designed to transmit something like 1000 times as much light -- a matter of considerable importance when such sharp bands are used, even in solar studies.

IV. References

1) B. Lyot, Comptes Rendus 197, 1593 (1933)

2) Y. Ohman, Nature 141, No. 3560, 157 (1938)

Nature 141, No. 3563, 291 (1938)

Populär Astronomisk Tidskrift, No. 1-2, pp. 11 and 27 (1938)

3) B. Lyot, Ann. Astrophys. 7, No. 1-2, p. 31 (1944)

4) T. Preston, Theory of Light, 3rd Edition, MacMillan, p. 397 (1901)

5) J. Evans, Ciencia e Investigación (Argentina) Vol. III, No. 9, p. 365 (1947)

6) B. Billings J.O.S.A., 37, p. 738 (1947)