LMSC-D674593

(NASA-CR-158702) RESEARCH ON SPECTROSCOPIC IMAGING. VOLUME 1: TECHNICAL DISCUSSION (Lockheed Missiles and Space Co.) 207 p HC A10/MF A01 CSCL 20F

N79-25872

Unclas G3/74 15377

RESEARCH ON SPECTROSCOPIC IMAGING

VOLUME I
TECHNICAL DISCUSSION

APRIL 1979

Prepared by

ALAN TITLE

WILLIAM ROSENBERG

LOCKHEED PALO ALTO RESEARCH LABORATORIES

3251 Hanover Street

Palo Alto, California 94304

Prepared for

LUNAR AND PLANETARY PROGRAMS

NASA HEADQUARTERS

Washington, D. C. 20546

UNDER CONTRACT NASW-3107

LMSC-D674593

RESEARCH ON SPECTROSCOPIC IMAGING

VOLUME I
TECHNICAL DISCUSSION

APRIL 1979

Prepared by
ALAN TITLE
WILLIAM ROSENBERG

LOCKHEED PALO ALTO RESEARCH LABORATORIES
3251 Hanover Street
Palo Alto, California 94304

Prepared for
LUNAR AND PLANETARY PROGRAMS
NASA HEADQUARTERS
Washington, D. C. 20546

Contents

(1) Introduction

The purpose of this report is to introduce the reader to the principles of operation and capabilities of birefringent filter systems. Birefringent filters have a number of significant advantages over more conventional spectral filtering techniques based upon gratings, Michelson interferometers, Fabry-Perot interferometric filters, or all deposited interference filters. Their primary advantage is the wide field of view compared to the other interference filters. The small angle sensitivity yields throughput gains of between 50 and 300 for imaging or collecting light from extended sources. This means a factor of 50 to 300 in number of resolution elements or total light collected in the desired spectral band. Besides the throughput advantage, birefringent filters have the capability of being tuned over 1.25 decades in frequency, can have tailored spectral profiles, and can have a set of selectable bandwidths. The devices are virtually insensitive to the direction of gravity. Furthermore, the filters can easily operate from $0^{o}C$ to $50^{o}C$ without any active temperature control.

Birefringent filters have long been used by solar physicists to study spectral characteristics of the solar atmosphere, mostly in the Hydrogen alpha line. Their use by scientists in other disciplines, however, has been very limited for a number of reasons. Commercial filters have been normally available only for a single wavelength, although two commercial tunable filters have been made. The filters are expensive, with single-frequency, 25 mm clear aperture filters costing about $50,000. Perhaps most importantly, the operation of the filters has not been understood by potential users and the basic sources of information about the filters have required knowledge of wave propagation of polarized light in isotropic media at a discouragingly high level. In addition, the basic work on birefringent filters has been published in foreign language journals for which only private translations are available.

We feel that the major limitation on the use of birefringent filters has been the lack of a single reference where an interested scientist can obtain an overall picture of the principles of operation and capabilities of these systems. This report is written to provide this overall view.

Normally, birefringent filters are explained using the techniques for analyzing propagation of polarized light, namely, the Jones Calculus. In this report the Jones techniques are avoided because of their unfamiliarity to the general reader and, perhaps more importantly, because of the lack of insight they provide into the behavior of birefringent plates and polarizers. While the Jones Calculus does an excellent job of determining the action of polarizing systems, it often does not make it physically clear why the optical system operates as it does. Fortunately, birefringent filters can also be explained by means of pulse response techniques, which only require some bookkeeping and a limited familiarity with the ideas of Fourier analysis. The result of this approach is a better physical understanding, but less mathematical analysis capability.

Chapter 2 is devoted to explaining the three fundamental types of birefringent filters: the Lyot (perfect internal polarizers), the Solc (no internal polarizers), and the filter with partial internal polarizers. The most straightforward of the filters, the Lyot, is used to introduce the basic concepts and analysis methods for birefringent systems. The filter with no internal polarizers is discussed next. Here it is expected that the reader will have to make some effort to follow the somewhat tedious, but reasonably coherent explanation. Chapter 2 ends with a discussion of partial polarizing filters which, while mathematically more complex than Solc filters, represent a small logical step from the Solc.

Chapter 3 is concerned with how the bandwidth or full width at half maximum (FWHM), free spectral range (FSR), finesse ($\mathscr{F}$), and profile shape are determined for birefringent filters. The main point of the chapter is to give the reader the tools to determine the basic design parameters of a filter system. Explored in some detail is the subject of profile shape which is a parameter which is not usually under the control of the filter designer, whereas for birefringent filters, virtually any profile shape can be achieved.

Chapter 4 discusses the subject of filter tuning. The methods for tuning the three basic filter types are established and some design implications of various tuning techniques are discussed. It is assumed

in the tuning section that achromatic tuning components are available. Specifically, it is the ready availability of easily fabricated achromatic wave plates which makes the methods discussed in this section possible.

The birefringent widefield element is introduced in Chapter 5. While a thorough analysis requires a large amount of tedious algebra, a simplified summary is presented as briefly as possible. The details of the field of view effects are deferred to an appendix, while only the important results and design considerations are presented in Chapter 5.

Chapter 6 introduces the subject of bandpass control. Here it is demonstrated that it is possible to construct Lyot and partial polarizing filters which can have a preselected set of bandwidths. Solc filters are not discussed because freedom to adjust all plates means that the bandpass of a Solc filter can be continuously adjusted.

The engineering considerations involved in reducing the theory to a workable filter system are discussed in Chapter 7. While this chapter does not describe all the considerations well enough to enable a novice to build birefringent filters, it does indicate potential difficulties and their solutions. Since birefringent filters are sophisticated optical instruments they require some care in their fabrication. Chapter 7 will indicate, however, that their tolerance requirements can be much less severe than other, comparably powerful, optical systems.

Chapter 8 provides a concise summary of the capabilities of birefringent filters. It highlights their operating principles and advantages over other filter systems. In addition, some of the current work in extending the range of applicability of these filters is outlined.

Finally, two appendices are included which discuss in more detail some topics which are felt to be important. Appendix 1 is a detailed analysis of the field of view properties of birefringent filters. Although the results of this analysis are summarized in Chapter 5, this is such an important aspect of these filters that its inclusion is beneficial. The engineering details, specifically for space qualifiable filters, are discussed in Appendix 2. There, a more thorough presentation is given than was included in Chapter 7.

Throughout this report, spectral transmission characteristics are always plotted for a full free spectral range of the shortest element in the filter. The plots are always for transmission with respect to phase retardation in the shortest element, which corresponds directly to the frequency of the incident radiation.

This report, then, should yield understanding of the physical properties and control techniques for birefringent filters. With this knowledge the reader should be able to specify but not design a filter. For those desiring more detailed knowledge of the details of filter operation and analysis, we have prepared a companion volume to this report that represents a reasonably complete set of scientific papers on birefringent filters. It includes translations of most of the original Lyot and Solc work, and reprints of many of the important recent papers. Space only permitted the reproduction of papers specifically devoted to filter techniques. For background material on polarized light and the Jones Calculus techniques, the excellent summary by Shurcliff and the reprint collection edited by Swindell are recommended.

(2-A) Lyot Filters (Perfect Internal Polarizers)

Birefringent filters were introduced by Lyot in 1933, and also developed independently by Ohman in 1938. These filters have been described in the literature, particularly in the papers by Lyot (1944) and Evans (1949) which are reproduced in Volume II of this report. The basic theory is repeated here because of its simplicity and as an introduction to the pulse response technique of birefringent filter analysis.

A Lyot filter is made up of a sequence of elements, each of which consists of a birefringent plate between perfect polarizers. As will be true of all birefringent plates in this discussion, the plate is assumed to be a uniaxial crystal cut with its optic axis in the plane of its faces. For a Lyot element the plate is placed with its optic axis oriented at 45^{o} to the orientation of the parallel entrance and exit polarizers. When viewed through a spectrograph the transmission of such an element appears as a series of alternate light and dark bands, a channel spectrum. Specifically, the transmission, τ, of the element is:

$$\tau(\lambda) = \cos^2\left(\frac{\pi \Delta d}{\lambda}\right) \tag{1}$$

with

$$\Delta = (n_o - n_e) \ , \tag{2}$$

where d is the thickness of the plate, Δ is the difference between the ordinary, n_o, and extraordinary, n_e, indices of refraction, and λ is the wavelength. A thick plate produces a channel spectrum of many thin bands closely spaced while a thin plate produces a channel spectrum of fewer thick bands more widely separated.

A Lyot filter, in its simplest form is made of a cascade of n such elements, each twice the thickness of its predecessor. The important feature of a cascade of Lyot elements is that the perfect polarizers isolate the elements so that the total transmission function is simply the product of the transmission functions of all the elements. Thus, if the elements have thicknesses

$$d_k = 2^k \ d_o \qquad k = 0, \ n - 1 \ , \tag{3}$$

the total transmission function becomes

$$\tau(\lambda) = \prod_{k=o}^{n-1} \cos^2\left(\frac{\pi 2^k d_o \Delta}{\lambda}\right). \tag{4}$$

For maximum transmission at an operating wavelength, λ_o, the total retardation, d_o, of the thinnest plate should be an integral number of wavelengths. The thickness, d_o is chosen to satisfy

$$d_o = \frac{m\lambda_o}{\Delta} \tag{5}$$

for some integer, m. All thicker plates will automatically satisfy the same criterion.

Shown in Figure 1 is the transmission function of each of the elements of a four-element Lyot filter, together with the cumulative transmission from each subfilter of the cascade.* These are shown in decreasing order of thickness of the elements. It is clear from the figure that the action of each successively thinner element is to double the separation between transmission maxima while changing the shape of each maximum only slightly.

The separation between transmission maxima is the free spectral range (FSR) of the filter. As is evident from Figure 1, the FSR of the overall filter coincides with the FSR of the thinnest element. Substituting (5) into (1) we see that at the operating wavelength, λ_o, the transmission function is

$$\tau(\lambda_o) = \cos^2\left(\frac{\pi\Delta d_o}{\lambda_o}\right) = \cos^2(m\pi) = 1. \tag{6}$$

The next transmission maximum will occur at a wavelength

$$\lambda_1 = \lambda_o + \lambda_{fsr},$$

*Remember that all transmission functions are with respect to phase retardation in the thinnest element, which corresponds directly to frequency.

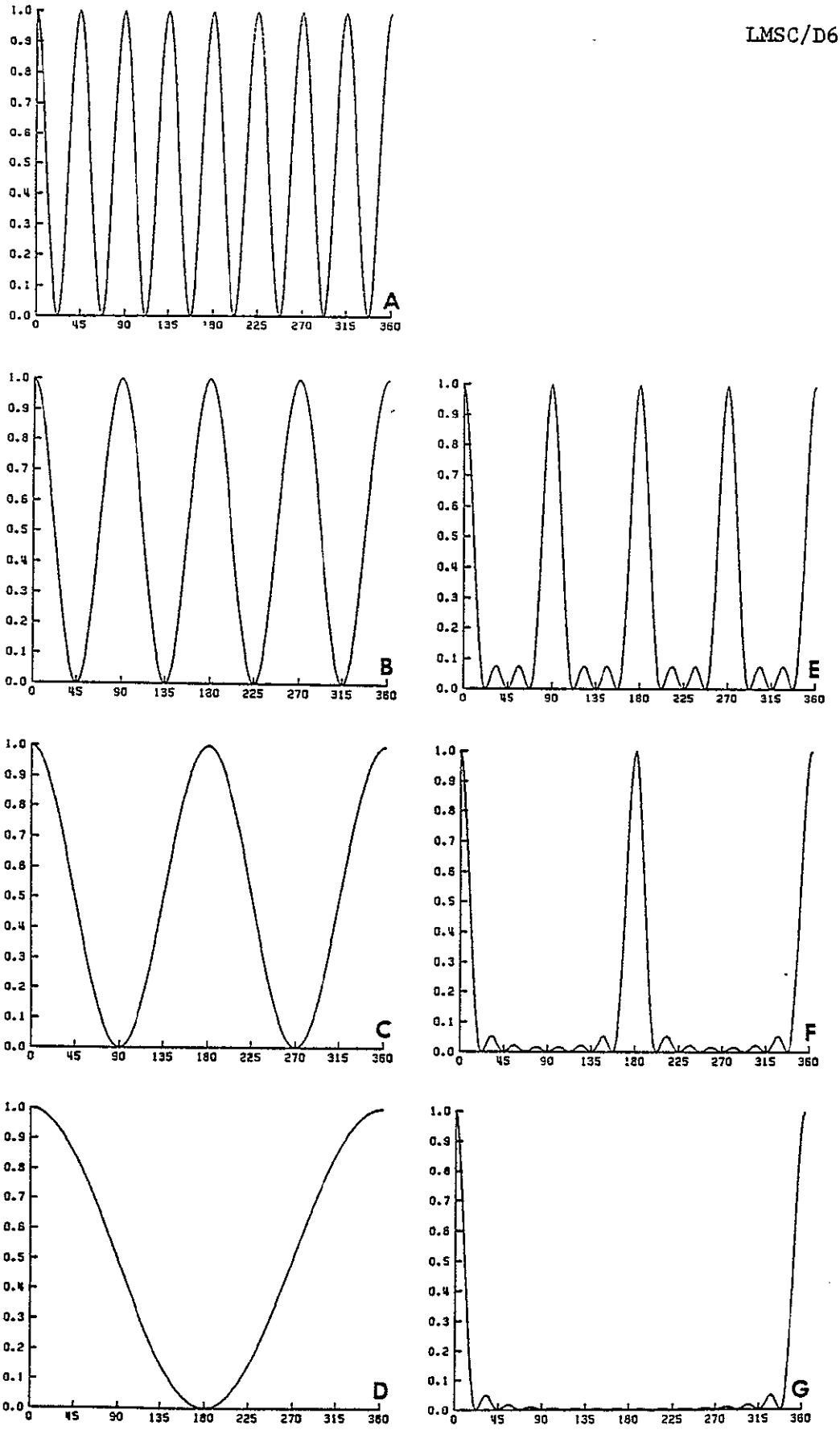


Figure 1. Transmission versus frequency for the components of a Lyot filter (a,b,c,d) and the transmission from each stage of the cascade (e,f,g)

such that

$$\frac{\pi \Delta d_o}{\lambda_o + \lambda_{fsr}} = (m-1)\pi . \tag{7}$$

In usual practice

$$\lambda_o >> \lambda_{fsr}$$

and (7) may be approximated by:

$$\frac{\pi \Delta d_o}{\lambda_o} (1 - \frac{\lambda_{fsr}}{\lambda_o}) = (m-1)\pi ,$$

which implies that the free spectral range, λ_{fsr}, becomes

$$\lambda_{fsr} = \frac{\lambda_o^2}{\Delta d_o} . \tag{8}$$

If it is assumed that the profile of the transmission maximum at λ_o is the same as that given by the thickest element then the filter's bandpass may be calculated. If the half width at half maximum of the transmission profile of the thickest element is λ_{hw}, then (1), (3), and (5) imply:

$$\cos^2\left(\frac{\pi 2^{n-1} \Delta d_o}{\lambda_o + \lambda_{hw}}\right) = \frac{1}{2} \tag{9}$$

and that

$$\frac{n-1 \Delta d_o}{\lambda_o + \lambda_{hw}} = m - \frac{1}{4} . \tag{10}$$

Again assuming that $\lambda_o >> \lambda_{hw}$

$$\lambda_{hw} = \frac{\lambda_o^2}{2^{n+1} \Delta d_o} \tag{11}$$

The filter bandwidth, or more precisely, the full width at half maximum (FWHM) is

$$\text{FWHM} = \frac{\lambda_o^2}{2^n \Delta d_o} . \tag{12}$$

In the subsequent analysis this estimate will be refined to account for the entire filter instead of just the thickest element, but (12) is a good approximation. Using (8) and (12), the finesse, $\mathscr{F}$, of a Lyot filter may be approximated as

$$\mathscr{F} = \frac{\text{FSR}}{\text{FWHM}} = 2^n . \tag{13}$$

In practice the best Fabry-Perot interferometric filters with apertures greater than a few millimeters have a finesse of 20 to 30. This is comparable to a 4 to 5 element Lyot cascade.

As mentioned in the introduction, a technique which has proven to be especially useful in optical filter analysis is impulse response or time domain analysis. This is a technique used regularly by electrical engineers and geophysicists, but only recently by optical designers. Impulse response analysis is complementary to the more common spectral analysis, although the two are mathematically equivalent. Either the time domain or the frequency domain response is sufficient to characterize a system since the two are conjugates under the Fourier transform operation. The freedom to investigate a filter's behavior in either domain is often quite useful. In particular, much of the experience of electrical engineers in time domain analysis is directly applicable to birefringent filter analysis and synthesis. Surprisingly, impulse response analysis of birefringent filters is actually much simpler than impulse response of more general electrical networks.

In Figure 2 the impulse response of a single Lyot element is diagrammed. A unit amplitude impulse (indicated by an arrow) emerging from the entrance polarizer has two components, one along and one orthogonal to the optic axis of the plate. These each have amplitude $1/\sqrt{2}$. Light propagates in these two polarizations with velocities c/n_o and c/n_e respectively. As a

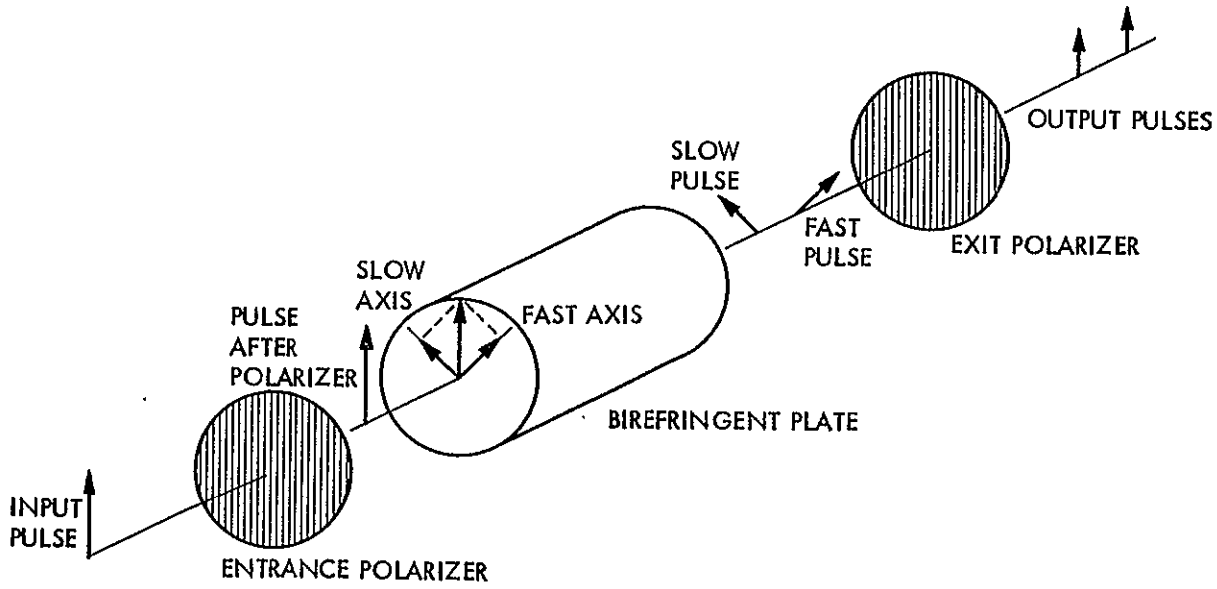


Figure 2. Pulse propagation through a simple Lyot element.

result the two component impulses incident on the crystal emerge at different times with a separation, Δt, given by

$$\Delta t = \frac{d}{c} (n_o - n_e). \tag{14}$$

When these two impulses encounter the exit polarizer, only their components in the pass polarization orientation are transmitted. The transmitted components again have relative amplitudes $1/\sqrt{2}$. The response of the entire element to the single incident unit impulse is therefore, a pair of pulses of amplitude 1/2 separated in time by Δt as given by (14). A thick element produces pulse pairs with long separations while a thin element produces closely spaced pairs.

When two Lyot elements, whose length ratio is 2:1, are cascaded, the output from the first element serves as input to the second. One input pulse creates four output pulses. More generally, the n element cascade discussed above will have 2^n output pulses for each input pulse. In Figure 3 the pulse response for a four-element Lyot filter has been illustrated. There, as is often done for convenience, the time axis has been redefined such that the output pulses are symmetric about the time origin. Half the output thus leads the input while the other half lags. This is a mathematical convenience and has no physical significance. It is clear from Figure 3 that the factor of 2 length reduction between successive crystals is the property which creates equally spaced, identical output pulses at each stage of the cascade.

Some properties of the impulse response are evident from Figure 3. The spacing of the pulses is

$$\Delta t_o = \frac{d_o}{c} (n_o - n_e) \tag{15}$$

and is determined by the thinnest element. The total time duration, T, of the response is

$$T = (2^n - 1)\Delta t_o. \tag{16}$$

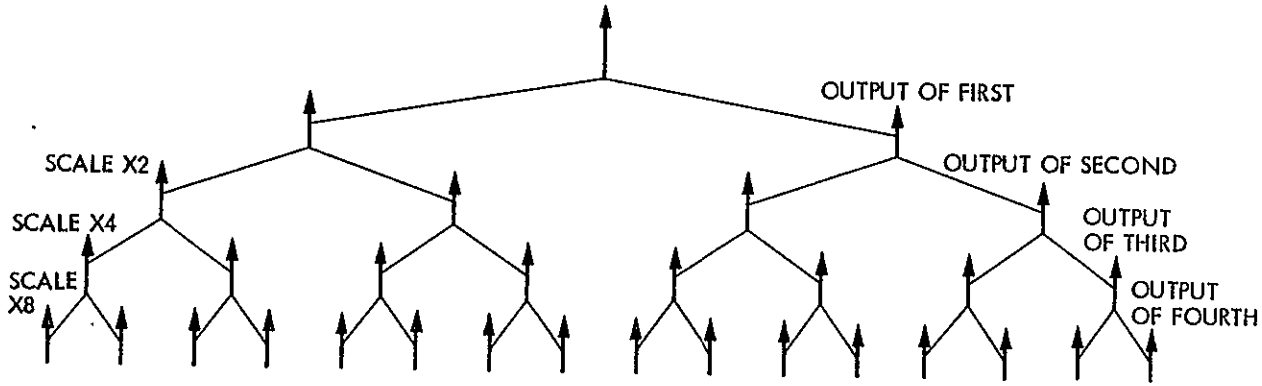


Figure 3. Pulse tree of a four-element Lyot filter. The pulse output of each element of the cascade is shown.

When n is moderately large so that

$$2^n >> 1$$

Then T is approximately

$$T \simeq 2t_n \tag{17}$$

where t_n is the time delay of the longest element.

In both the spectral analysis illustrated in Figure 1, and the pulse response shown in Figure 3, it was assumed that the longest element was first, followed by successively shorter elements. It should be evident that the spectral response, being the product of single element spectral responses, is unchanged by a rearrangement of elements. Equally true, however, though perhaps not obvious, is the invariance of the pulse response to rearrangements. The detailed structure of the pulse response tree of Figure 3 changes, but the final pulse output will remain unchanged. It is a useful exercise to verify this for at least one rearrangement.

As mentioned previously, the pulse response and the spectral response are intimately related to each other through the Fourier transform operation. Since we have already derived the two responses for Lyot filters, we can demonstrate this connection. Some mathematical notation and definitions will be needed and are summarized below.

An impulse at time t_k will be represented by the Dirac delta function, $\delta(t - t_k)$, which has the properties:

$$\delta(t-t_k) = 0 \quad \text{for} \quad t \neq t_k$$

$$\int_{-\infty}^{\infty} f(t)\delta(t-t_k)dt = f(t_k) \; . \tag{18}$$

The Fourier transform, $F(\nu)$, of a function, $f(t)$, is defined as:

$$F(\nu) = \int_{-\infty}^{\infty} f(t)\ e^{i2\pi\nu t} dt \tag{19a}$$

with the inverse transform being:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\nu)\ e^{-i2\pi\nu t} d\nu\ . \tag{19b}$$

Any specific properties of the Fourier transform will be introduced as needed.

Using the above definitions, the equivalence of the pulse and spectral responses of a single Lyot element may be demonstrated easily. Assume that the pulse response of Figure 2 is represented by

$$f(t) = \frac{1}{2}\,\delta(t + \frac{\Delta t}{2}) + \frac{1}{2}\,\delta(t - \frac{\Delta t}{2}). \tag{20}$$

This represents a pair of pulses separated by Δt as given by (14) The Fourier transform of (20) is

$$\begin{aligned} F(\nu) &= \frac{1}{2} \int_{-\infty}^{\infty} \delta(t + \frac{\Delta t}{2})\ e^{-i2\pi\nu t} dt + \frac{1}{2} \int_{-\infty}^{\infty} \delta(t - \frac{\Delta t}{2})\ e^{-i2\pi\nu t} dt \\ &= \frac{1}{2}\, e^{-i\pi\nu\Delta t} + \frac{1}{2}\, e^{i\pi\nu\Delta t} \\ &= \cos\pi\nu\Delta t\ . \end{aligned} \tag{21}$$

$F(\nu)$ as given by (21) is the amplitude response of the element. The transmission is given by the squared magnitude of the amplitude response which becomes

$$\tau(\lambda) = \cos^2 \frac{\pi d(n_o - n_e)}{\lambda} \tag{22}$$

when we substitute, $\nu = \frac{c}{\lambda}$, and, $\Delta t = \frac{d}{c}(n_o - n_e)$. This is the same transmission function as given previously, (1), for a single Lyot element.

For a two-element Lyot filter the transform of the pulse response is:

$$F(\nu) = \frac{1}{4} \int_{-\infty}^{\infty} \delta(t + \tfrac{3}{2}\Delta t) + \delta(t + \tfrac{1}{2}\Delta t) + \delta(t - \tfrac{1}{2}\Delta t) + \delta(t - \tfrac{3}{2}\Delta t) \; e^{i2\pi\nu t} dt$$

$$= \frac{1}{2} \cos 3\pi\nu\Delta t + \frac{1}{2} \cos \pi\nu\Delta t \tag{23}$$

where Δt is again given by (14), and is the time delay of the shorter element. Using the trigonometric identity

$$\cos A + \cos B = 2 \cos \left(\frac{A+B}{2}\right) \cos \left(\frac{A-B}{2}\right) \tag{24}$$

and the relation for ν the transmission becomes

$$\tau(\nu) = \cos^2 \frac{2\pi c\Delta t}{\lambda} \cos^2 \frac{\pi c\Delta t}{\lambda} . \tag{25}$$

The equations (21) and (23) suggest that for an n element Lyot cascade there will be a sum of 2^n impulses in the pulse response, a sum of 2^{n-1} cosine terms in the amplitude spectral response, and a product of n cosine squared factors in the transmission function. Though this may be proven by repeated use of identity (24), here only the result will be presented.

$$\frac{1}{2^{n-1}} \sum_{k=1}^{2^{n-1}} \cos \frac{(2k-1)\pi c\Delta t}{\lambda} = \prod_{k=o}^{n-1} \cos \frac{2^k \pi c\Delta t}{\lambda} . \tag{26}$$

It is not true in general that any sum of cosines can be written as a single product, but the relationship between these sum and product forms will be used in further arguments. The specific result, (26), for the Lyot filter is a consequence, again, of the 2:1 length ratio between successive elements.

Using the pulse response—spectral response relationship, it is easy to find a closed form expression for the product of cosines, (26). In this case it is convenient to redefine the time origin to coincide with the earliest output pulse, resulting in a transformed response:

$$F(\nu) = \frac{1}{2^n} \int_{-\infty}^{\infty} \sum_{k=o}^{2^n-1} \delta(t-k\Delta t)\ e^{i2\pi\nu t} dt$$

$$= \frac{1}{2^n} \sum_{k=o}^{2^n-1} e^{i2\pi k\nu\Delta t} \tag{27}$$

Since this is simply the sum of a geometric progression it has the closed form evaluation

$$F(\nu) = \frac{1}{2^n} \quad \frac{1 - e^{i2\pi 2^n \nu \Delta t}}{1 - e^{i2\pi\nu\Delta t}}$$

$$= \frac{1}{2^n} \quad \frac{\sin 2^n \pi\nu\Delta t}{\sin\pi\nu\Delta t} \quad \frac{e^{i\pi 2^n \nu\Delta t}}{e^{i\pi\nu\Delta t}} \tag{28}$$

As the transmission is simply the squared magnitude,

$$\tau(\lambda) = \frac{1}{4^n} \quad \frac{\sin^2(\frac{2^n \pi c\Delta t}{\lambda})}{\sin^2(\frac{\pi c\Delta t}{\lambda})} \; . \tag{29}$$

Using expression (29) it is possible to re-derive some of our earlier results, but with more accuracy. In particular, it is evident that unit intensity transmission maxima occur whenever $\frac{c\Delta t}{\lambda}$ is an integer. This is the same result as (6). Now, however, it may be seen that the half maximum transmission occurs whenever

$$\frac{1}{2} \cong \frac{1}{4^n} \quad \frac{\sin^2(2^n \pi \frac{c\Delta t \lambda_{hw}}{\lambda_o^2})}{\sin^2(\frac{c\Delta t \lambda_{hw}}{\lambda_o^2})} \; . \tag{31}$$

When $2^n >> 1$ it is possible to simplify (31) further, yielding

$$\sin\,(2^n \pi \frac{c\Delta t \lambda_{hw}}{\lambda_o^2}) = \frac{2^n}{\sqrt{2}} \quad \frac{c\Delta t \lambda_{hw}}{\lambda_o^2} \tag{32}$$

This is a transcendental equation of the form

$$\sin q = \frac{1}{\sqrt{2}}\, q \tag{33}$$

whose solution is

$$q = 1.39 \tag{34}$$

This implies that the full width at half maximum is

$$\text{FWHM} = .44 \frac{\lambda_o^2}{2^{n-1} c\Delta t} \tag{35}$$

which is a more accurate result than the earlier estimate, (12), based on the length of the longest element alone.

The free spectral range is, as given previously in (8),

$$\lambda_{fsr} = \frac{\lambda_o^2}{c\Delta t}$$

Consequently the finesse is now given as

$$\mathscr{F} = 1.13\ 2^n \tag{36}$$

The subject of Lyot filters and their modifications will be returned to in later chapters.

(2-B) Solc Filters (No Internal Polarizers)

The Lyot filter discussed previously was invented in 1933 and was developed into a practical solar instrument during the period 1933 to 1950. In 1953 the next major advance in birefringent filter design came when Solc realized that a filter could be made without internal polarizers. The explanation of the basic operation of the Solc filter first appeared in English in 1958 by Evans. By 1964 Harris and his associates had developed a general procedure for the synthesis of arbitrary spectral response profiles with Solc-type filters. In this report all such filters without internal polarizers will continue to be called Solc or Solc-type filters even though their development owes much to Harris and others whose contributions will be described subsequently. As will be seen, the principle of operation of a Solc filter is far less intuitive than that of a Lyot filter. This is reflected in the literature on the subject prior to 1967, none of which provided a reasonably clear explanation of the mechanism. Fredga and Hogbom (1967) introduced the concept of pulse paths and path amplitudes which provides a simple qualitative explanation of the Solc filter and a reasonably close quantitative approximation for most Solc designs. Volume II of this report contains reprints of translations of the original Solc papers from the Czechoslovakian Journal of Physics, in addition to a Solc paper which appeared in JOSA, the Evans paper mentioned earlier, the Harris synthesis procedure, and the Fredga and Hogbom paper.

The original Solc filters, introduced in 1953, were made as a cascade of identical birefringent plates, at specified orientations, between entrance and exit polarizers. In this discussion, the angular orientations will be denoted as:

ϕ_o - orientation of entrance polarizer (usually taken to be zero)

ϕ_k - orientation of the fast axis of the k-th plate

ϕ_p - orientation of exit polarizer

Solc presented two different types of filters which, while equivalent, are not identical. These are known as the fan and folded filters and are specified as:

$$\text{Fan filter} - \begin{aligned} \phi_o &= 0 \\ \phi_k &= (2k-1)\alpha \qquad 1 \le k \le n \\ \phi_p &= 0 \end{aligned} \tag{1}$$

$$\text{Folded filter} - \begin{aligned} \phi_o &= 0 \\ \phi_k &= (-1)^k \alpha \qquad 1 \le k \le n \\ \phi_p &= 90^o \end{aligned} \tag{2}$$

where $$\alpha = 45^o/n \tag{3}$$

for both configurations.

In Figure 1 the transmission profiles for several fan type filters are shown along with the corresponding spectral responses of equal length Lyot filters. The figure demonstrates some interesting properties of fan filters. The first is that they are not dramatically different from the Lyot filters. This is a surprising feature to many who are accustomed to the Lyot design with internal polarizers. The second property is that for very few elements the sidelobe level is lower than the corresponding Lyot profile, but for moderate and large values of n the sidelobe level is higher. This behavior was noticed early in the study of these filters, and the design was criticized for its relatively poor sidelobe behavior. As will be shown later, the reason for this behavior is not too difficult to understand and the mechanism to improve the performance follows easily.

Transmission profiles for the folded filters are not presented since they have a remarkable equivalence to the fan filter profiles. A folded filter has the exact same transmission profile as a fan filter except that the spectrum is shifted by one half of the free spectral range. For this reason, fan type filters will be the principal subject of this discussion.

Evans (1958) presented a general formula for the transmission profile of the fan (and thus the folded) Solc filters. This transmission is:

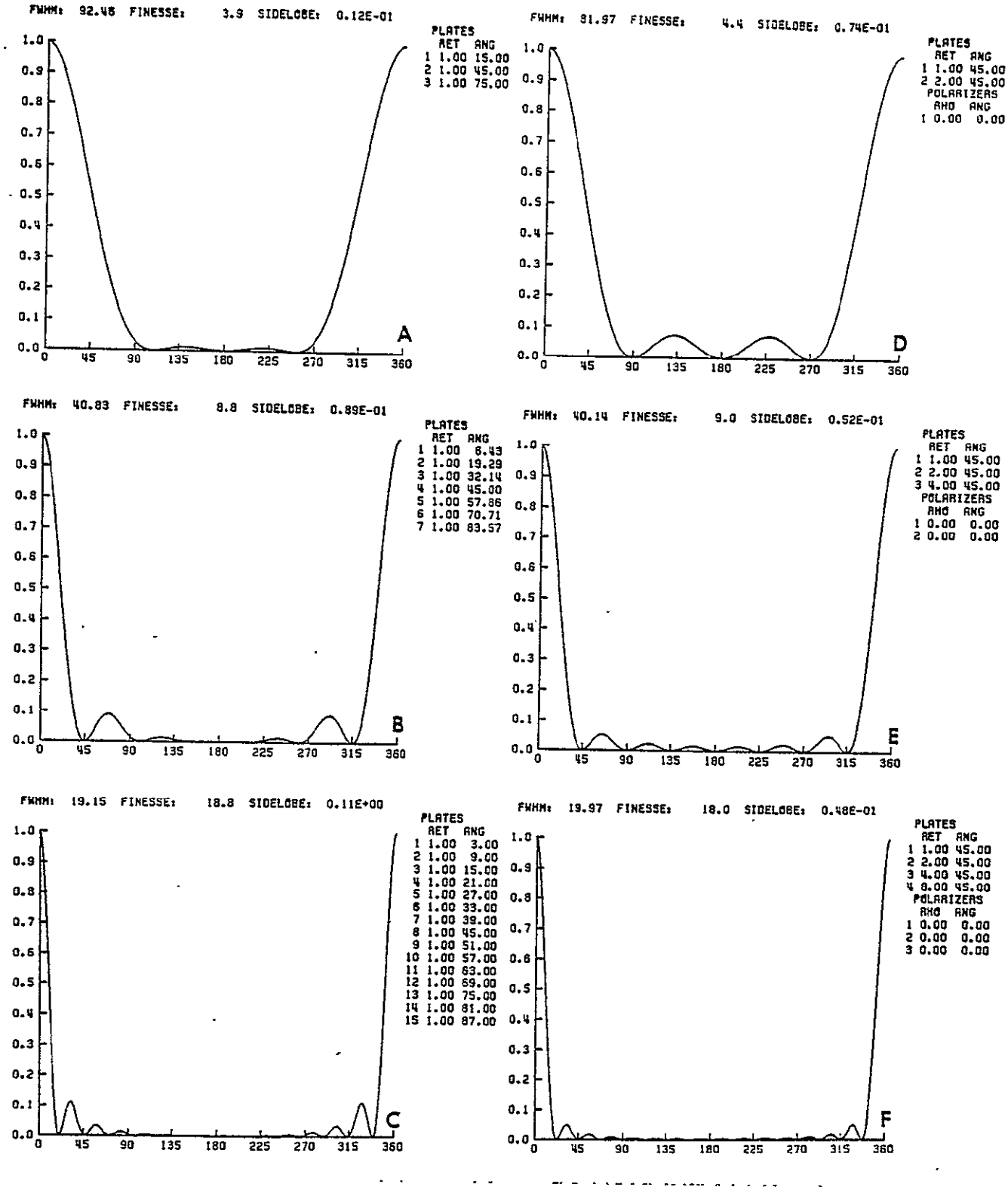


Figure 1. Transmission versus frequency for 3, 5, and 15 element fan Solc filters (a,b,c) and 2, 3, and 4 element Lyot filters (d,e,f).

$$\tau(\lambda) = \left[\frac{\sin n\chi}{\sin \chi} \cos \chi \tan \alpha\right]^2$$

where $$\cos \chi = \cos \alpha \cos \frac{\pi c \Delta t}{\lambda} \qquad (4)$$

and Δt is the time delay of a single birefringent plate as in the Lyot filter analysis. Transmission maxima occur, as stated earlier, at wavelengths for which $c\Delta t = k\lambda$. Formula (4) was derived using a powerful technique known as the Jones Calculus (R. C. Jones, (1941)), but, while elegant, the derivation gave us little insight into the mechanism of the Solc filter.

The lack of an intuitive explanation was typical of the literature on Solc filters. While the Jones Calculus was a powerful analytic tool it did not indicate the enormous potential of the Solc filters.

Solc (1959) showed that by a slight modification of the angles given by (1)-(3), the sidelobe level could be reduced by a factor of 5 or more. This provided an answer to the earlier criticism of the filters, but only touched on the real capability, namely that by a suitable modification it is possible to achieve almost any spectral profile.

As examples of the capabilities of Solc filters, three responses have been selected. The first response, shown in Figure 2(a) is the transmission profile of a seven element Solc filter which is equivalent to a three-element Lyot filter. Both have identical, uniform pulse responses and, consequently, identical transmission profiles. Figure 2(b) is the transmission profile of a filter whose pulse response has a linear taper, or triangular shape. With increasing time delay the impulse response rises linearly to a central maximum and then decays linearly back down again. This results in a further sidelobe reduction over the uniform pulse response. If the pulse response amplitudes are made proportional to binomial coefficients, $\binom{n}{k}$, a transmission profile with no sidelobes at all results. This profile is shown in Figure 2(c). In order to facilitate comparisons, the orientations and selected profile characteristics are reproduced here:

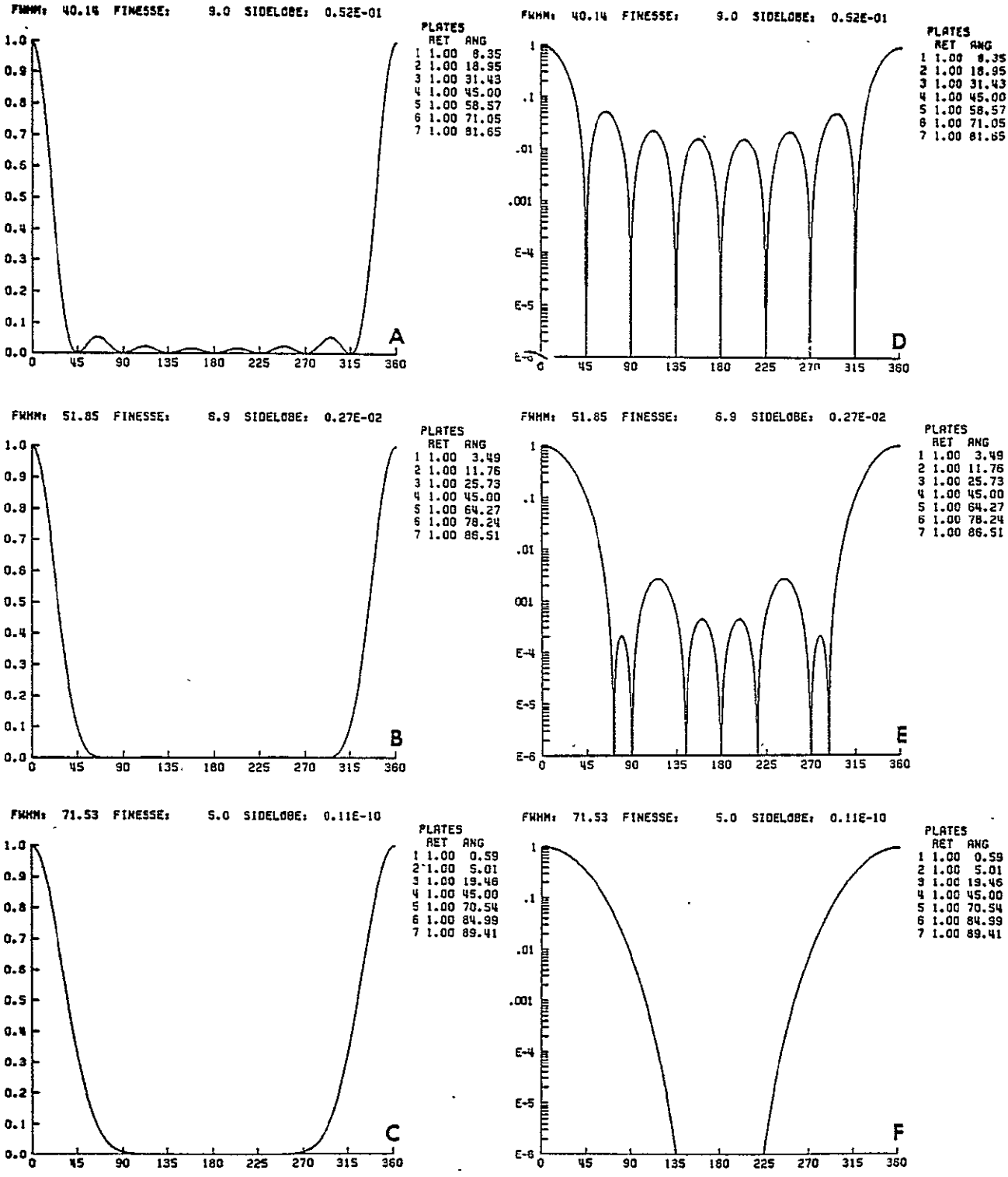


Figure 2. Transmission versus frequency for Solc filters with uniform (a), linear taper (b), and binomial pulse responses (c) on linear (a,b,c) and logarithmic scales (d,e,f).

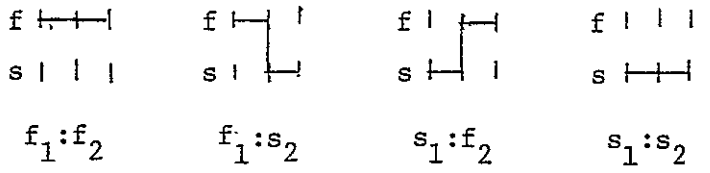
	Basic Solc	Lyot	Triangle	Binomial
ϕ_1	6.4°	8.4°	3.5°	0.6°
ϕ_2	19.3°	18.9°	11.8°	5.0°
ϕ_3	32.1°	31.4°	25.7°	19.5°
ϕ_4	45.0°	45.0°	45.0°	45.0°
ϕ_5	57.9°	58.6°	64.3°	70.5°
ϕ_6	70.7°	71.1°	78.2°	85.0°
ϕ_7	83.6°	81.6°	86.5°	89.4°
FWHM/FSR	.113	.112	.144	.199
Sidelobe	.089	.052	.003	0.0

As the figure and the table show there is no limit to the sidelobe reduction which is possible with suitable angular modifications from the basic Solc design. The implication is that with the appropriate procedure for determining the plate orientations, any transmission profile is available. In the following, two techniques for determining plate angles corresponding to a given response will be described.

An elementary description of the Solc filter operation was provided by Fredga and Hogbom (1967). Here, for the first time, the concept of pulse paths and path amplitudes was introduced. Harris et al (1964) had previously used pulse response techniques in their synthesis procedure, but not in an intuitively appealing development. Unfortunately, the simplicity of the pulse response analysis of Lyot filters doesn't quite transfer to Solc filters. The difficulties arise because the plates are not isolated from each other by polarizers as they are in the Lyot configuration. This means that at the interface between two plates the impulse transmission from one plate to the other must be considered in two orthogonal polarization states instead of just one. Mathematically this means that the impulse response must be considered as a vector quantity instead of just a scalar.

In this section each of the two polarization components will be considered separately. Any pulse incident on a plate will be decomposed

into two pulses, one along the fast axis and one along the slow axis of the plate. This is the natural coordinate system for the plate and is the key to an understanding of the entire system. The logical consequence of this is that a pulse incident on the first plate produces two output pulses. These in turn result in four pulses output from the second plate, eight from the third, and finally 2^n from the last plate and from the exit polarizer. They, of course, occur at only n+1 distinct time delays so there is considerable overlap. It is much easier, however, to calculate the amplitude of each of the 2^n pulses individually than to try to calculate the pulse response directly. This is the path amplitude approach, since each of the 2^n output pulses can be imagined to be the result of a specific path through the system. Each pulse will have followed a path through some fast axes and some slow axes. The paths through a two-plate filter can be diagrammed as:



(5)

The advantage of this approach is the ease of calculating the pulse amplitude associated with each path. This is illustrated by a two-plate example.

In Figure 3, a general Solc configuration with two plates between entrance and exit polarizers is illustrated. The respective orientations are ϕ_o, ϕ_1, ϕ_2, ϕ_p. In the analysis of such a system the absolute orientations are unnecessary since the filter must be invariant to overall rotations. The quantities of interest are the angle differences which will be denoted:

$$\theta_k = \phi_k - \phi_{k-1} \qquad 1 \leq k \leq n$$
$$\theta_p = \phi_p - \phi_n \,. \tag{6}$$

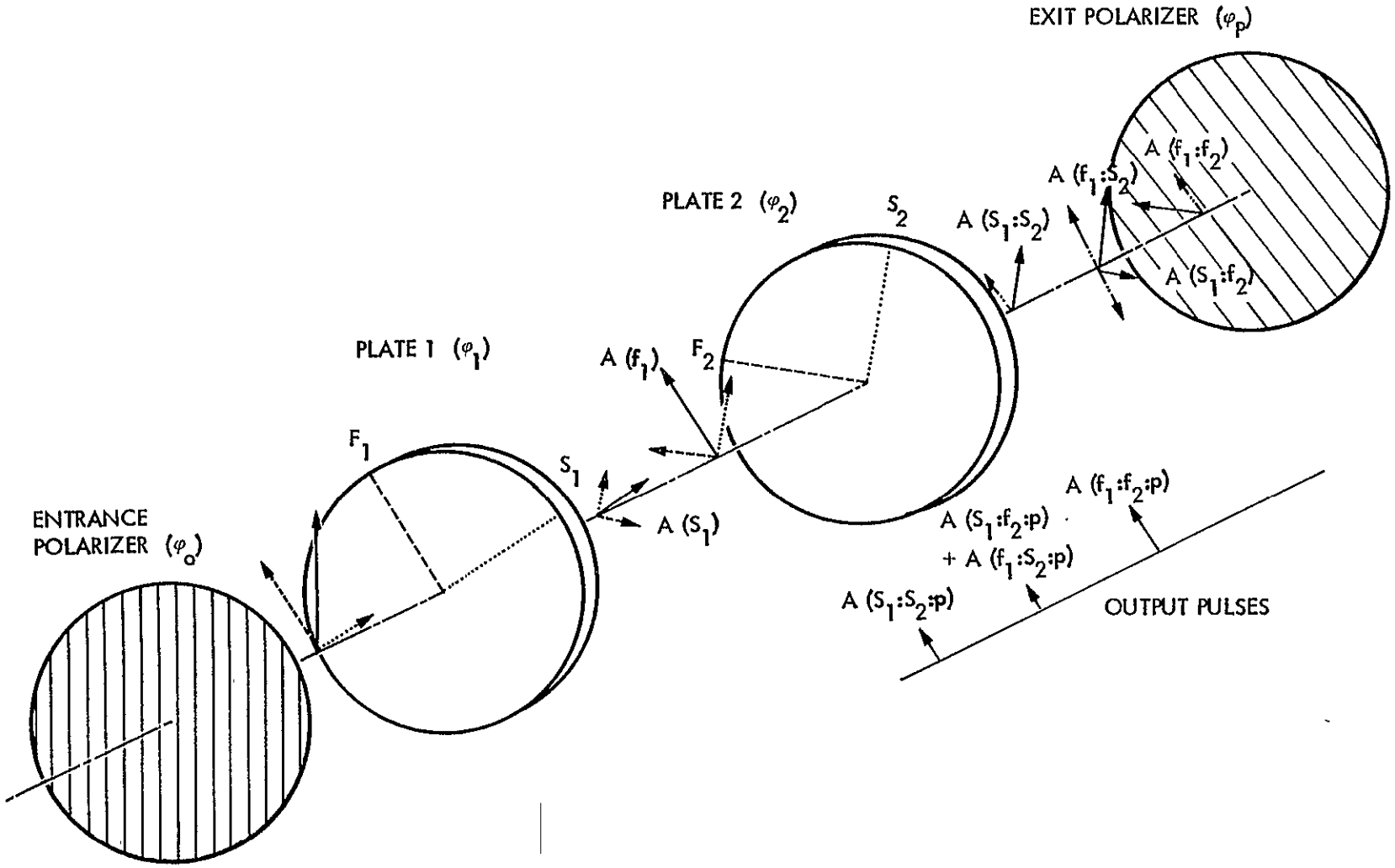


Figure 3. Pulse propagation through a general two-element Solc filter.

With this notation, the response of the filter to an incident impulse may be considered. Such an impulse will generate two components incident on the first birefringent plate; one along the fast axis direction and one along the slow axis direction. These components will have relative amplitudes:

$$\begin{aligned} A(f_1) &= \cos \Theta_1 \\ A(s_1) &= \sin \Theta_1 \end{aligned} \tag{7}$$

respectively. Each of the two pulse outputs of the first plate may be similarly decomposed into components along the fast and slow axes of the second plate and the resulting amplitudes calculated. These become:

$$\begin{aligned} A(f_1{:}f_2) &= \cos \Theta_2 \; A(f_1) = \cos \Theta_2 \cos \Theta_1 \\ A(s_1{:}f_2) &= -\sin \Theta_2 \; A(s_1) = -\sin \Theta_2 \sin \Theta_1 \\ A(f_1{:}s_2) &= \sin \Theta_2 \; A(f_1) = \sin \Theta_2 \cos \Theta_1 \\ A(s_1{:}s_2) &= \cos \Theta_2 \; A(s_1) = \cos \Theta_2 \sin \Theta_1 . \end{aligned} \tag{8}$$

Finally the exit polarizer at orientation, ϕ_p, will transmit impulse components whose amplitudes are:

$$\begin{aligned} A(f_1{:}f_2{:}p) &= \cos \Theta_p \; A(f_1{:}f_2) = \cos \Theta_p \cos \Theta_2 \cos \Theta_1 \\ A(s_1{:}f_2{:}p) &= \cos \Theta_p \; A(s_1{:}f_2) = -\cos \Theta_p \sin \Theta_2 \sin \Theta_1 \\ A(f_1{:}s_2{:}p) &= -\sin \Theta_p \; A(f_1{:}s_2) = -\sin \Theta_p \sin \Theta_2 \cos \Theta_1 \\ A(s_1{:}s_2{:}p) &= -\sin \Theta_p \; A(s_1{:}s_2) = -\sin \Theta_p \cos \Theta_2 \sin \Theta_1 . \end{aligned} \tag{9}$$

The notation used in (7) - (9), the colon notation, indicates the local orientation of a scalar impulse component as it goes through each plate. This is the pulse path concept introduced by Fredga and Hogbom in which the output impulse response is decomposed into an enumeration of all possible paths through the plate system. With each path is associated an amplitude, as in (9), and a time of arrival. The time of arrival depends on the number of slow axes traversed by the path since each slow axis traversal results in a relative delay, Δt, with respect to a fast axis traversal. In keeping with the convention that the impulse response of a system is centered at the time origin, the output

time for any path will be determined by the simple rule:

Rule 1: Each fast axis traversal contributes a time delay of $-\Delta t/2$
Each slow axis traversal contributes a time delay of $+\Delta t/2$.

In addition, the notation, $D(t_o)$, will be used to mean $\delta(t-t_o)$, an impulse at time t_o. The total impulse response for a general two-plate system becomes, therefore:

$$\begin{aligned} A_2(t) = &\cos\theta_1 \cos\theta_2 \cos\theta_p \, D(-\Delta t) \\ &-[\sin\theta_1 \sin\theta_2 \cos\theta_p \\ &+ \cos\theta_1 \sin\theta_2 \sin\theta_p] \, D(0) \\ &- \sin\theta_1 \cos\theta_2 \sin\theta_p \, D(+\Delta t) . \end{aligned} \tag{10}$$

Since the spectral response depends only on the relative amplitudes of the pulse response, it is convenient to normalize (10) such that the earliest output pulse has unit amplitude. The response is then:

$$\begin{aligned} A_2(t) = &D(-\Delta t) \\ &- [\tan\theta_1 \tan\theta_2 \\ &+ \tan\theta_2 \tan\theta_p] \, D(0) \\ &- \tan\theta_1 \tan\theta_p \, D(+\Delta t) . \end{aligned} \tag{11}$$

It should be apparent from (11) and the preceding development that the pulse response may be derived by some simple rules for evaluating path amplitudes simply from the plate to plate transitions of each path. The form of these rules will be taken to be the normalized form, to produce the amplitudes (11) rather than (10). The specific rule is:

Rule 2: The first plate contributes a factor 1 to the fast path and a factor $\tan\theta_1$ to the slow path
Each subsequent transition contributes a factor according to:

$f_{i-1}:f_1$	1
$f_{i-1}:s_i$	$\tan \Theta_i$
$s_{i-1}:f_i$	$-\tan \Theta_i$
$s_{i-1}:s_i$	1

The exit polarizer contributes a factor according to:

$f_n:p$	1
$s_n:p$	$-\tan \Theta_p$.

Finally, the normalization constant is:

$$\cos \Theta_p \prod_1^n \cos \Theta_i .$$

Using these rules the total pulse response may be computed by a brute force enumeration procedure. Though this is tedious, the pulse response may often be approximated without a total enumeration, and as a theoretical tool, the path enumeration procedure has been very valuable. Many general properties of Solc filters are proven most easily by arguments concerning individual path amplitudes. As the enumeration is conceptual rather than computational in those cases, the efficiency is not an issue.

For an n element Solc filter there are 2^n possible paths but only n+1 possible time delays. In order to understand the pulse response from the path enumeration it is necessary to divide the set of paths into subsets of equal time delay.

The path which only traverses fast axes appears at the relative time, $-\frac{n}{2}\Delta t$. Similarly the all-slow path emerges at time, $+\frac{n}{2}\Delta t$. These are, of course, unique. The output pulse at time, $-\frac{n-2}{2}\Delta t$, will be the resultant of the paths which include exactly one slow axis. There are n of these since there are n plates. In general, the output pulse at relative time $-\frac{n-2k}{2}\Delta t$ $\quad k = 0,\ldots,n$ will be the resultant of $\binom{n}{k}$ different paths.

Now, if Rule 2 is considered, it is evident that the f:f and s:s transitions do not attenuate the relative amplitude, but that f:s and s:f transitions contribute a $\tan\Theta$ attenuation. For the usual case in which $\Theta < 45^\circ$ this is indeed an attenuation. This means, that for each time delay, those paths with the fewest axis type changes (f:s or s:f) will contribute the largest amplitudes to the resultant output pulse.

To first order, therefore, the paths which dominate the pulse response, in increasing order of time delay, are given as:

$$
\begin{aligned}
&A(f_1{:}f_2{:}\ldots{:}f_n{:}p) = 1 && t = -\frac{n}{2}\Delta t \\
&A(f_1{:}f_2{:}\ldots{:}f_{n-1}{:}s_n{:}p) \\
&\qquad = -\tan\theta_n \tan\theta_p && t = -\frac{n-2}{2}\Delta t \\
&A(f_1{:}f_2{:}\ldots{:}f_{n-k}{:}s_{n-k+1}{:}s_n{:}p) \\
&\qquad = -\tan\theta_{n-k+1} \tan\theta_p && t = -\frac{n-2k}{2}\Delta t \\
&A(f_1{:}s_2{:}\ldots{:}s_n{:}p) = \tan\theta_2 \tan\theta_p && t = \frac{n-2}{2}\Delta t \\
&A(s_1{:}s_2{:}\ldots{:}s_n{:}p) = -\tan\theta_1 \tan\theta_p && t = \frac{n}{2}\Delta t\,. \qquad (12)
\end{aligned}
$$

A first order approximation to a filter design for a given pulse response would result from the n+1 uncoupled simultaneous equations (12). Take, for example, the construction of a seven-plate filter whose pulse response is equivalent to a three-element Lyot filter, that is, eight equal amplitude pulses. Equations (12) imply that the θ_i are all equal and that

$$1 = -\tan\theta_i \tan\theta_p\,. \qquad (13)$$

This is satisfied, and, since the normalization constant is maximized, so is the transmission, when

$$
\begin{aligned}
\theta_i &= \frac{90}{n+1} \\
&= 11^{\circ}.25 \qquad (14)
\end{aligned}
$$

and

$$
\begin{aligned}
\theta_p &= \theta_i - 90 \\
&= -78^{\circ}.75\,. \qquad (15)
\end{aligned}
$$

This is the filter whose entrance and exit polarizers are parallel at 0° orientation and the seven plates are uniformly spaced at angles

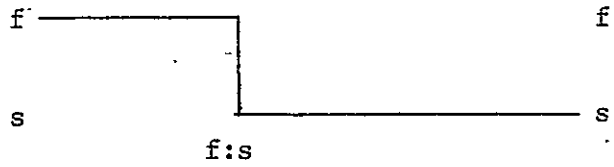
$$\phi_k = 11^{\circ}.25\ k \qquad k = 1,\ldots,7\,. \qquad (16)$$

Since the design criteria, equations (12), were only approximate it is not surprising that this filter only approximates a uniform amplitude pulse response. In particular the pulse response of this filter is:

Time (Δt)	$-\frac{7}{2}$	$-\frac{5}{2}$	$-\frac{3}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{5}{2}$	$\frac{7}{2}$
Amplitude	.17	.13	.11	.09	.09	.11	.13	.17

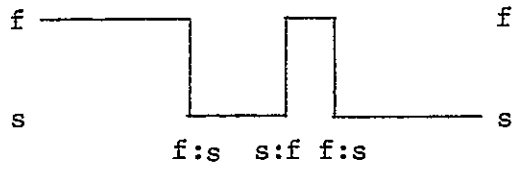
The spectral response of this filter is shown in Figure 4. As indicated there, the FWHM is 10.1% of the free spectral range, and the highest sidelobe is 13% of the peak transmission. This is contrasted with the spectral response of the three element Lyot filter shown in Figure 2(a). There the FWHM is 11.2% of the free spectral range but the sidelobe level is only 5.2% of the peak transmission.

A closer look at the impulse response of this first order approximation is quite revealing. The pulse amplitudes taper downward from the outside inward. This means that the central pulses are being attenuated much more than the outside pulses even though their dominant paths are all equal. The reason for this is that equation (12) has accounted for paths of the form



(17)

but not paths of the form



(18)

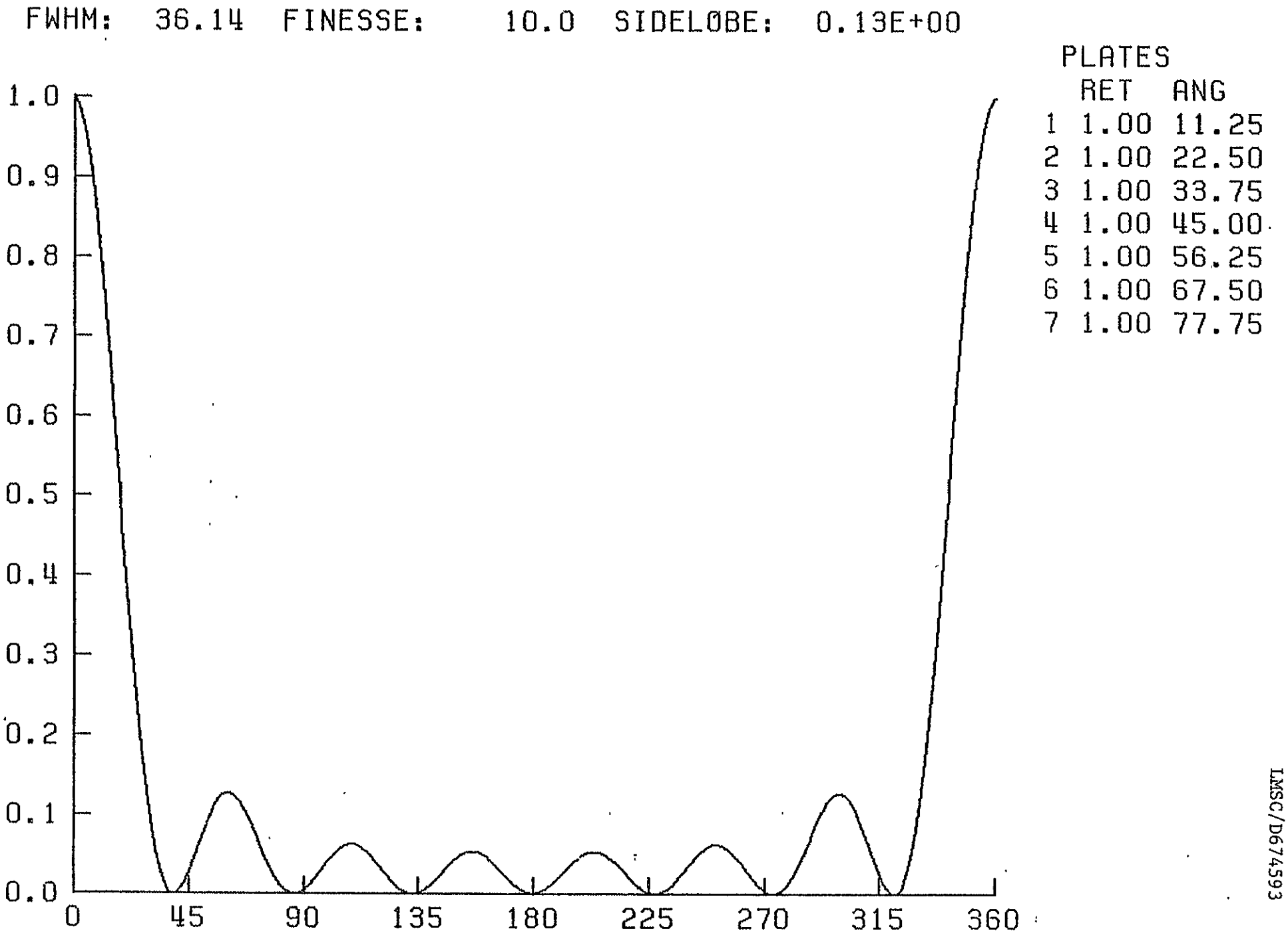


Figure 4. Transmission versus frequency for a seven-plate Solc filter with a uniform plate distribution.

since these paths are of third degree in $\tan\theta_i$ instead of first degree. There are, however, many such paths for the interior pulses (remember that there are $\binom{n}{k-1}$ paths for the k-th pulse) and they are all negative since they have one s:f transition. In particular there are

$$(n - k + 1)\ (k - 1) \tag{19}$$

paths of the form (18) for the k-th pulse. Since this is a quadratic function of k it gives the first order approximation its characteristic form. This pulse response is unfortunate since it is just the opposite of the response which minimizes sidelobes. Responses which minimize sidelobes peak in the center and taper downward towards the ends. In order to make a good filter, therefore, the angles must be modified to compensate for this effect. The Solc fan filter does this by making the first angle difference, θ_1, only half the others (while adjusting all the angles to maintain symmetry of course).

It is possible to incorporate the type (18) path amplitudes into a set of equations similar to (12) and solve for the higher order approximation to a given pulse response. This isn't too difficult, but it is a little too cumbersome to include here. Another, related, technique is to start with the first order approximation and to iterate to achieve a desired pulse response. This is possible since, for small angles, θ_i, the angular corrections, $\Delta\theta_i$, which need be made at each step are approximately linear functions of the pulse response error terms. Thus, at least for filters which have no extreme deviations from the basic Solc response, the Solc technique of making small angle modifications will converge to the desired filter. This is the result which Fredga and Hogbom present as a consequence of the path amplitude arguments they introduced: sidelobe reduction to any desired level may be achieved by small angle modifications from the basic Solc design, or, more generally, any n+1 term Fourier series approximation to an arbitrary transmission profile may be achieved by iteration from the first or second order filter configurations.

In 1964 Harris, Ammann, and Chang went further, but in a way which is less intuitive. They used some reasonably sophisticated mathematical

techniques to relate the impulse response to the roots of n-th degree polynomial equations. One result of this is a rigorous mathematical proof that any of the above mentioned transmission profiles can be realized by a Solc-type filter. Another result, however, is that for most transmission profiles the filter is not unique. There are essentially different plate angle configurations which give the same transmission profile. This result is non-intuitive and not widely realized, but has some practical implications for actual filter construction.

Consider a filter constructed to give a specified transmission profile, $|C(\lambda)|^2$. This filter will have an impulse response whose pulse amplitudes are:

$$C_i \qquad i = 0,\ldots,n \; . \tag{20}$$

If the exit polarizer were rotated 90° the filter would now have a transmission profile which is the complement of the original,

$$|D(\lambda)|^2 = 1 - |C(\lambda)|^2 \tag{21}$$

and an impulse response whose pulse amplitudes are denoted

$$D_i \qquad i = 0,\ldots,n \; . \tag{22}$$

The essence of the argument is that the square magnitude, $|C(\lambda)|^2$, or, $|D(\lambda)|^2$, does not uniquely specify the impulse responses, C_i or D_i, respectively. The filter is, however, specified by the impulse responses. Consequently any of the compatible C_i and D_i sequences will generate a different filter with the specified transmission profiles. In Figure 5(a), the transmission profile for a 9-element filter with a uniform pulse response is presented along with two different orientation configurations which achieve that profile. Notice that one of the angle sequences, ϕ_i, shows the monotonic progression typical of a modified Solc configuration. The other, however, has larger angle differences and a much less regular pattern. Both of these configurations imply the same transmission profile, $|C(\lambda)|^2$, rejection profile, $|D(\lambda)|^2$, and transmission impulse responce, C_i. They differ, however, in their rejection impulse responses, D_i. The

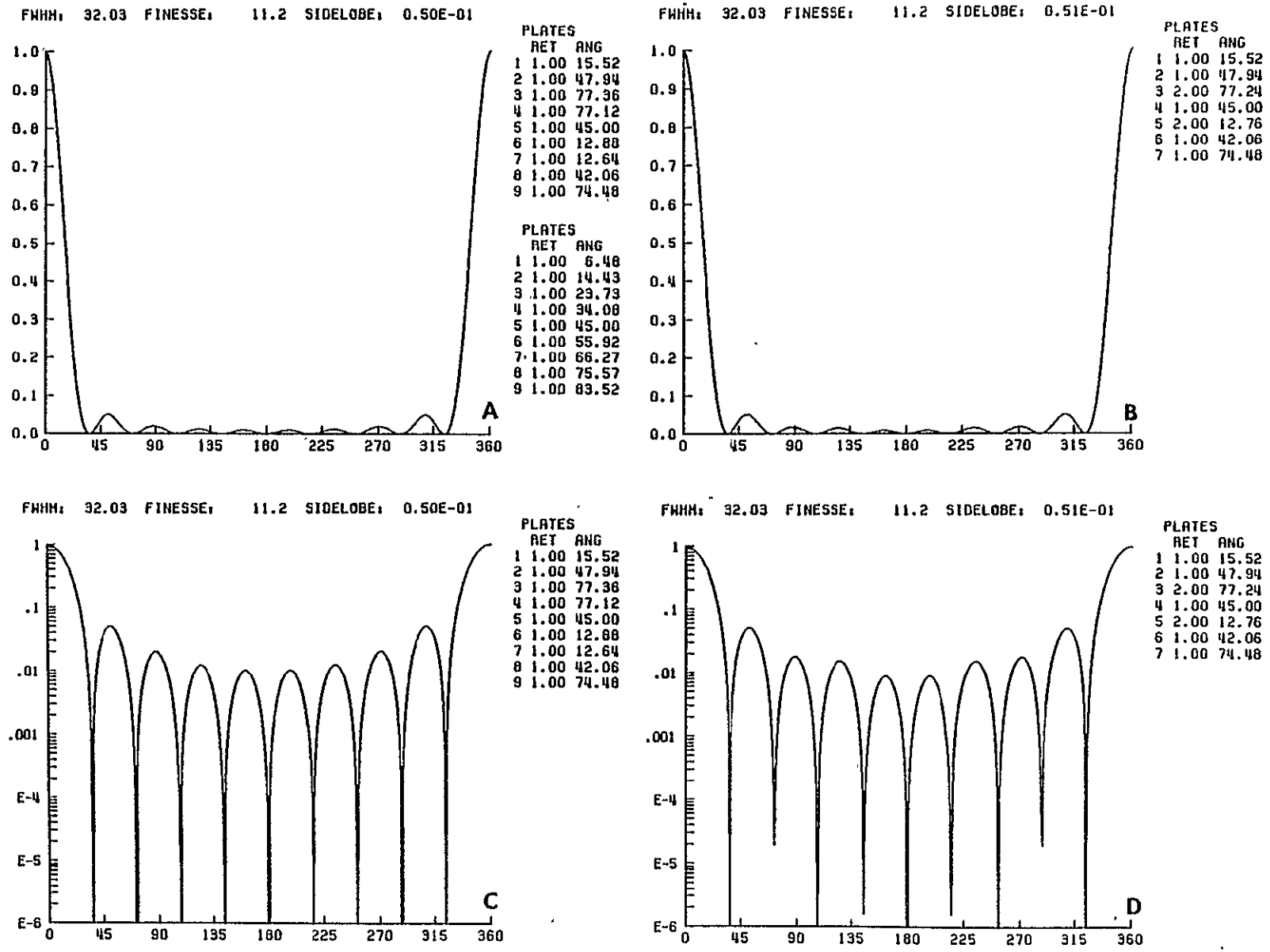


Figure 5. Transmission versus frequency of a nine-element Solc filter with a uniform pulse response (a) and its seven-element approximation (b) shown on linear (a,b) and logarithmic (c,d) scales.

impulse responses for these two filters are given as:

I	ϕ_i	6.5	14.4	23.7	34.1	45.0	55.9	66.3	75.6	83.5	
	C_i	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10
	D_i	-.01	-.03	-.04	-.06	-.09	-.11	-.14	-.18	-.22	.88
II	ϕ_i	15.5	47.9	77.4	77.1	45.	12.9	12.6	42.1	74.5	
	C_i	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10
	D_i	-.03	-.10	-.19	-.26	-.18	-.09	-.43	-.40	-.53	.36

The existence of these multiple solutions has some practical importance. Some of the solutions may turn out to be more tolerant of alignment errors than others, and this question is a current research topic. Another important question raised is whether any of the solutions have consecutive plates oriented at the same or nearly the same angle. This has the practical significance that a double length plate could be substituted for the two single length plates. This would save on fabrication complexity. In Figure 5(b), such a seven-element approximation to one of the nine-plate filters of Figure 5(a) is shown. Notice that two pairs of internal plates have been merged into double length plates at their average orientations. The performance is nearly identical.

The above example is not an unusual occurrence. In Figure 6(a) two filters which achieve the same, triangular, impulse response are presented. Again, there are plate pairs which are very nearly parallel and may be combined as shown in Figure 6(b). Again, the performance is nearly identical, with the seven-plate filter actually doing a little better.

Another important aspect of the Harris synthesis procedure is its value as a diagnostic tool. Since it determines the plate angles for a given transmission profile it may be used to adjust a filter. If the filter's transmission profile is measured, the synthesis algorithm may be used to determine the angle at which each plate is currently set. It is a simple matter, then, to readjust the angles to their desired settings.

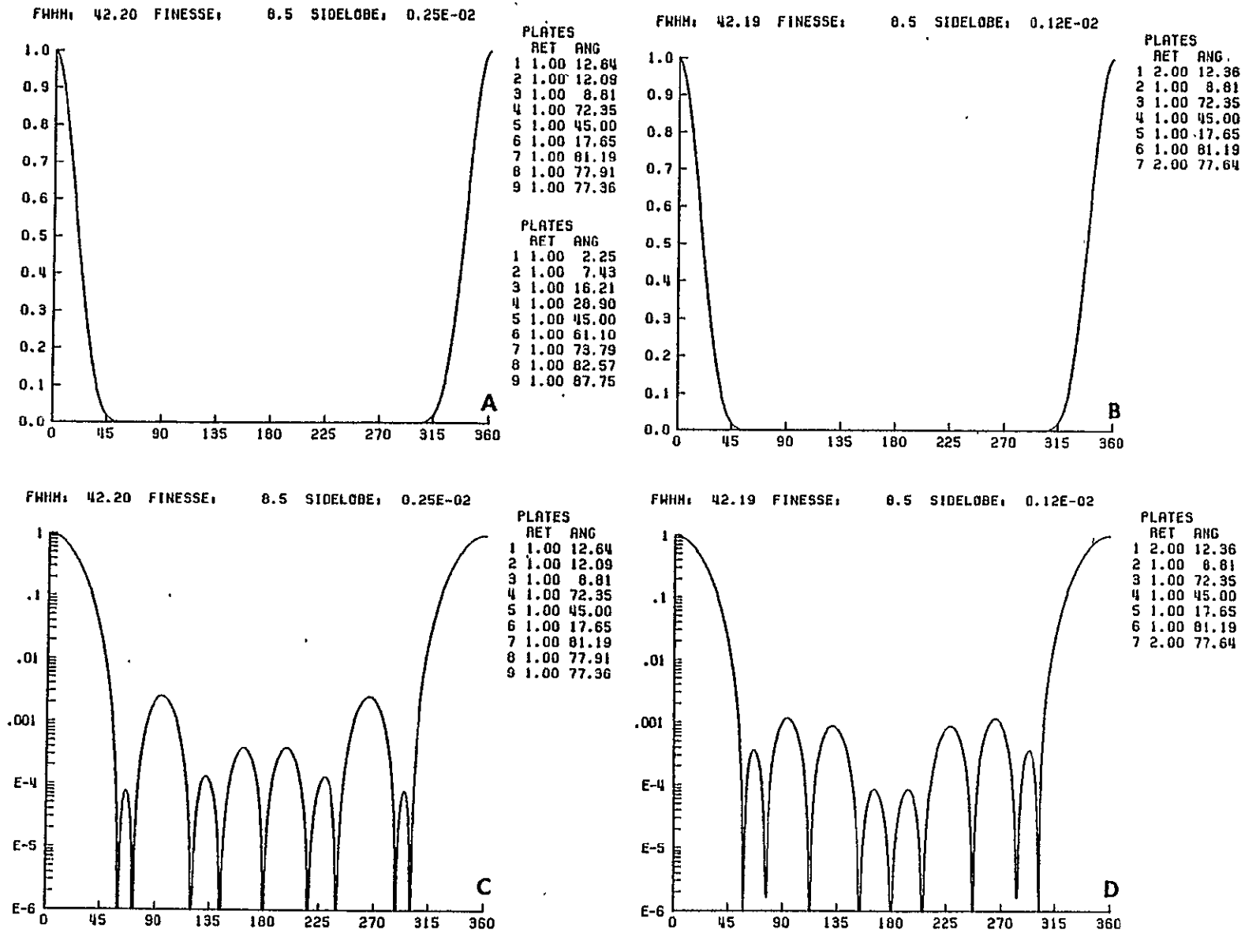


Figure 6. Transmission versus frequency of a nine-element Solc with a linear taper pulse response (a) and its seven-element approximation (b) shown on linear (a,b) and logarithmic scales (c,d).

(2-C) Partial Polarizer Filter

The effects of using imperfect, or partial, polarizers in birefringent filters were discovered in 1953, the same year as the introduction of the Solc filter. Giovannelli and Jefferies discovered in an analysis of Lyot filters that in some cases the secondary maxima were suppressed if the internal polarizers were imperfect (the ratio of transmissions in the rejection and pass directions was non-zero). In 1965 Beckers and Dunn verified the partial polarizer result via a computer program for Lyot filters. However, experimental attempts to verify the effect failed and the result was generally considered a curiosity. In 1974 Title gave a simple theoretical explanation of why a single partial polarizer worked in a Lyot filter, and later developed several filter designs based upon multiple partial polarizers. A number of filters were constructed using partial polarizers and the theoretical predictions were verified. Previous experiments were shown to have failed because of neglect of the birefringence of the sheet partial polarizer.

A partial polarizer differs from a perfect polarizer in that it not only has transmission in the pass direction, say the X direction, but also transmits in the orthogonal, Y, direction, Mathematically, the partial polarizer has two transmission amplitude coefficients, ρ_x and ρ_y. Usually ρ_x is near unity and ρ_y is much less. The key to understanding the action of a partial polarizer is to consider it equivalent to the parallel combination of two optical paths. The first path will have a perfect polarizer with transmission amplitudes

$$\rho_x' = \rho_x - \rho_y$$
$$\rho_y' = 0 . \qquad (1)$$

The second path will have a neutral density filter with a transmission amplitude

$$\rho_n' = \rho_y \qquad (2)$$

along all directions.

A general description of the effect of partial polarizers follows directly from the pulse path enumeration technique developed in the previous section. Each partial polarizer will have an easily describable effect on each of the pulse paths. Before developing the general analytic technique, however, it is useful to consider an important special case; the two-element Lyot filter with its intermediate polarizer replaced by a partial polarizer.

In Figure 1 the two-element Lyot filter with partial polarizer is shown in its conceptually equivalent form. It is depicted as a pure two-element Lyot in parallel with a pair of plates separated by a neutral density filter.

In a normal Lyot filter the plates may be at either plus or minus 45° with respect to the polarizers. Since the perfect polarizers isolate each element from its neighbors the relative orientation of successive plates does not affect the filter. When the polarizers are partial, however, the angles between the plates become important and must be considered in the analysis. This means that for the two-element example two cases must be distinguished; the plates are either parallel or perpendicular.

In the path of Figure 1 which contains the neutral density filter the time delay will be the sum of the time delays if the plates are parallel but will be the difference if the plates are perpendicular. The impulse response is, therefore,

$$A(\underline{\underline{I}}) = \frac{\rho_x - \rho_y}{4}\left[D(-\frac{3\Delta t}{2}) + D(-\frac{\Delta t}{2}) + D(\frac{\Delta t}{2}) + D(\frac{3\Delta t}{2})\right]$$

$$+ \frac{\rho_y}{2}\left[D(-\frac{(2\pm1)}{2}\Delta t) + D(+\frac{(2\pm1)}{2}\Delta t)\right] \qquad (3)$$

where the sum and difference in the second term corresponds to parallel and perpendicular respectively. The transmission of the filter is then,

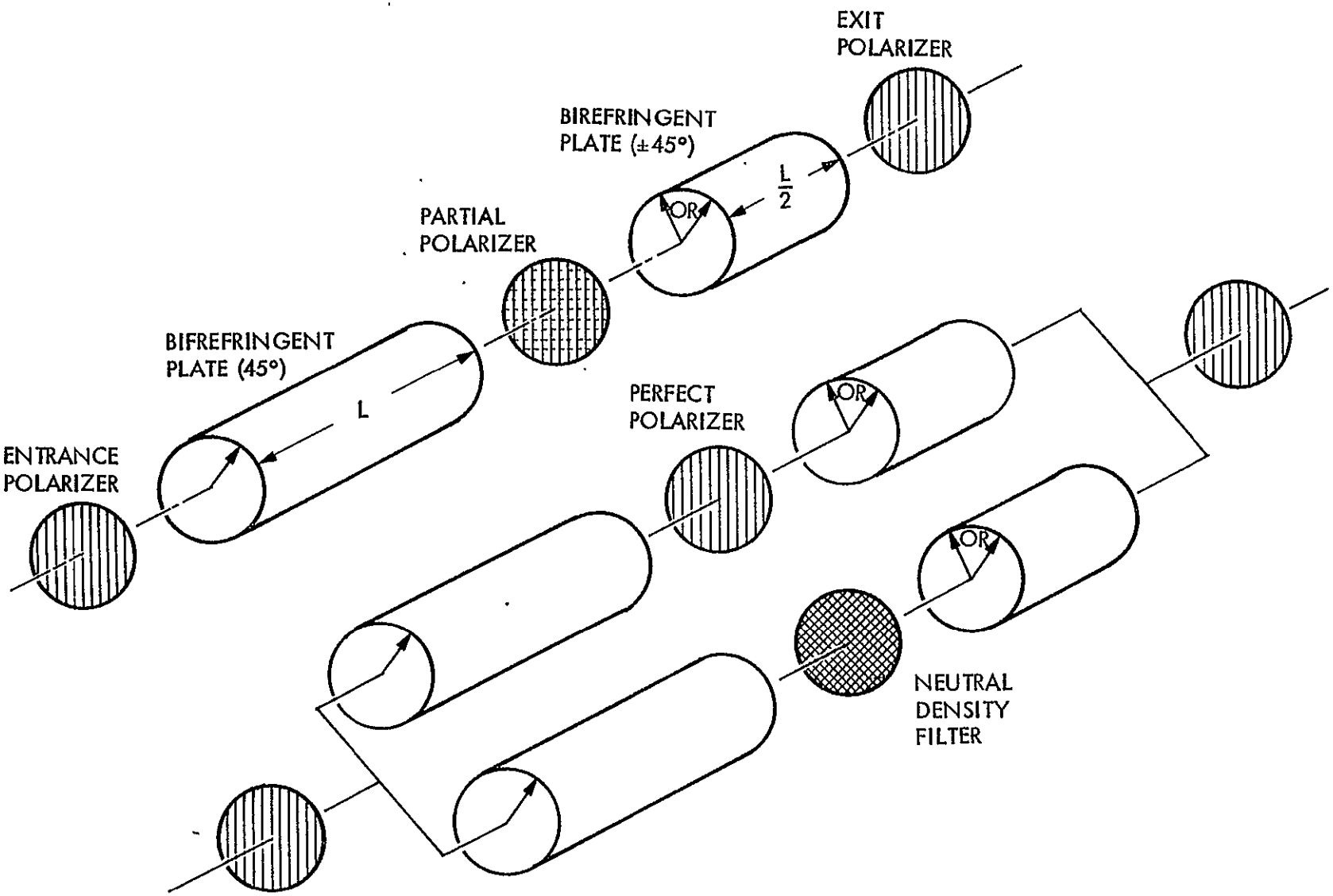


Figure 1. Optical schematic of a two-element partial polarizing filter and its conceptually equivalent form of a parallel two-element Lyot and two elements with a neutral density filter between them.

$$\tau\left(\begin{matrix}\parallel\\ \perp\end{matrix}\right) = \left[(\rho_x - \rho_y)\cos\left(\frac{2\pi c\Delta t}{\lambda}\right)\cos\left(\frac{\pi c\Delta t}{\lambda}\right) + \rho_y \cos\left(\frac{(2\pm1)\pi c\Delta t}{\lambda}\right)\right]^2. \quad (4)$$

Although (3) shows the distinction between the perfect polarizer and neutral density filter components of the pulse response, it is not a form which is very convenient for analysis. If, however, (3) is rewritten,

$$A_{\parallel}(t) = \left(\frac{\rho_x-\rho_y}{4}\right)\left[(1+2S)\ D(-\tfrac{3}{2}\Delta t) + D(-\tfrac{1}{2}\Delta t) + D(\tfrac{1}{2}\Delta t) + (1+2S)\ D(\tfrac{3}{2}\Delta t)\right] \quad (5)$$

$$A_{\perp}(t) = \left(\frac{\rho_x-\rho_y}{4}\right)\left[D(-\tfrac{3}{2}\Delta t) + (1+2S)\ D(-\tfrac{1}{2}\Delta t) + (1+2S)\ D(\tfrac{1}{2}\Delta t) + D(\tfrac{3}{2}\Delta t)\right] \quad (6)$$

with $$S = \frac{\rho_y}{\rho_x-\rho_y}$$

it is obvious how the transmission will be related to the polarizers. Since ρ_y, the rejection direction transmission amplitude, is always less than ρ_x, the pass direction amplitude, S is always positive. The parallel plate pulse response, (5), has a taper which is high at the ends and low in the center. As in the previous section, this taper produces large sidelobes. The perpendicular plate response, (6), has a taper which produces lower sidelobes. Specifically it is higher in the center than at the ends.

In Figure 2 the transmission profiles for parallel and perpendicular configurations are shown for several values of S. As may be seen, the parallel plate transmission sidelobes increase with S as expected, while the perpendicular plate sidelobes decrease. In fact, when S equals one, the sidelobes disappear entirely. The pulse distribution is proportional

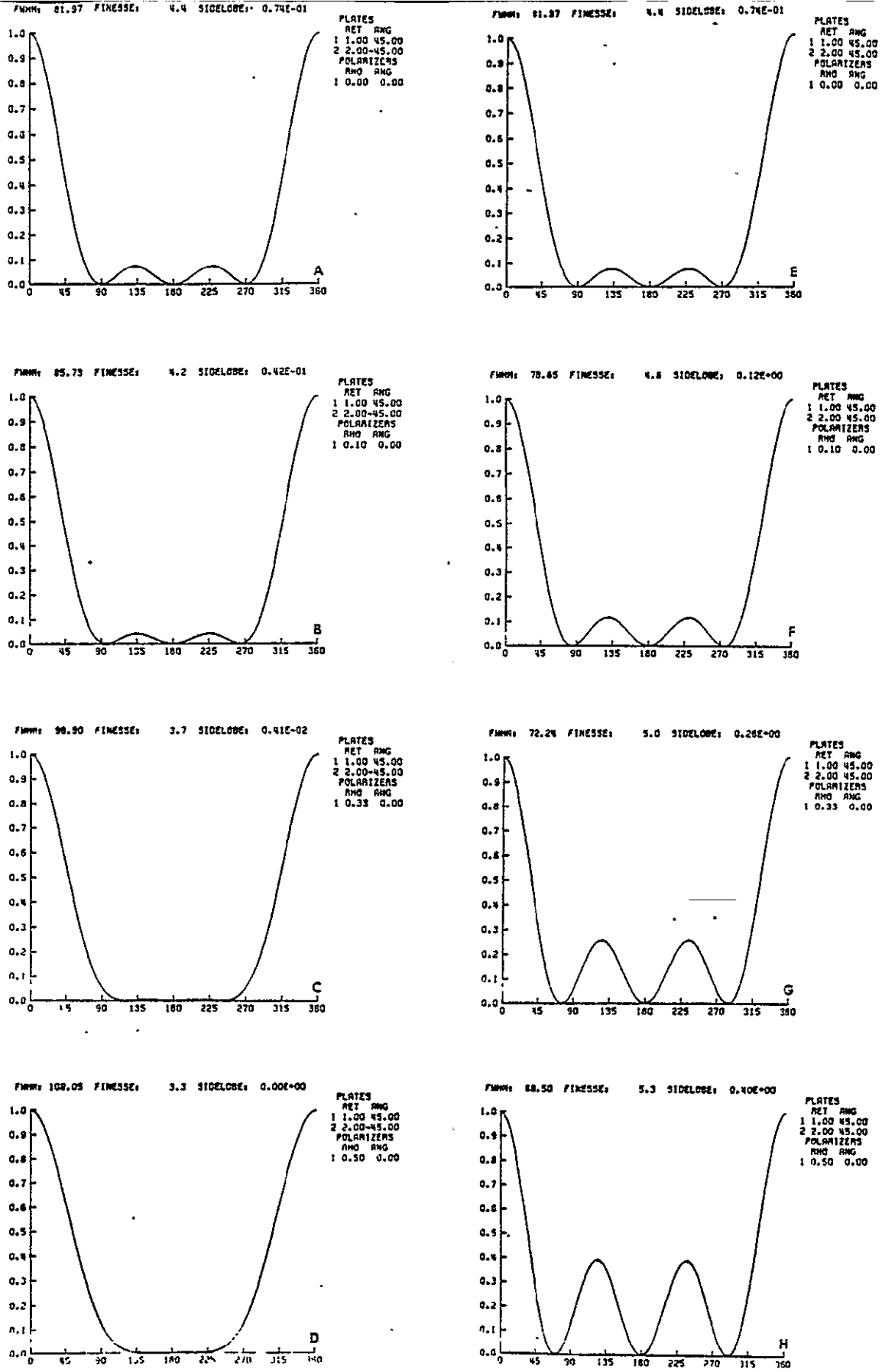


Figure 2. Transmission versus frequency for a partial polarizing module with crossed (a,b,c,d) and parallel plates (e,f,g,h) for S=0.0, 0.111, 0.5, and 1.0 (ρ=0.0, .1, .333, and .5).

to 1:3;3;1, a binomial distribution, and binomially distributed pulse responses have no sidelobes. For this minimum, the polarizer transmission amplitudes are in the ratio

$$\frac{\rho_y}{\rho_x} = \frac{1}{2} \tag{7}$$

and the transmission intensity ratio is

$$\frac{\tau_y}{\tau_x} = \frac{1}{4} \tag{8}$$

which is quite a poor polarizer.

A temptation, after seeing how much sidelobe control is available in a two-element Lyot configuration, is to replace more polarizers with partial polarizers in a multi-element Lyot. It is possible to use partial polarizers in a three-element Lyot filter and to achieve sidelobe reduction, but for more elements the use of partial polarizers raises the secondary maxima. There are designs, however, in which some of the polarizers are partial and some are perfect, $\rho_y = 0$, which are beneficial. Alternate partial polarizer configurations which are discussed in Title (1974), are very attractive.

The discussion above has been concerned with the special case of Lyot configurations. More generally it is possible to consider Solc configurations with the possibility of partial polarizers between some of the plates. Since it is known that the Solc configuration already has enough degrees of freedom to allow the synthesis of arbitrary transmission profiles, the addition of the polarizers would only increase the multiplicity of solutions for any given profile. This feature, however, might make it possible to find filter configurations which minimize the number of plates while still allowing for arbitrary profile control.

In order to analyze the general plate/partial polarizer configurations, the pulse path technique of the previous section has been useful. The partial polarizers introduce some more notation, but actually very little more complexity into the procedure. The notation introduced in the previous section is:

ϕ_o - orientation of entrance polarizer
ϕ_k - orientation of k-th plate $1 \leq k \leq n$
ϕ_p - orientation of exit polarizer

To this must be added:

ψ_k - orientation of the k-th partial polarizer (following k-th plate) $1 \leq k \leq n-1$

$Q_k - \left(\frac{\rho_x - \rho_y}{\rho_x + \rho_y}\right)$ for the k-th partial polarizer

Notice that a perfect polarizer is represented by Q = 1 while no polarizing is represented by Q = 0. For general partial polarizing, $0 \leq Q \leq 1$. With this notation, the rules for constructing pulse path amplitudes become:

$$A(f_k:p_k:f_{k+1}) = \cos(\phi_{k+1}-\phi_k) + Q_k \cos(\phi_k+\phi_{k+1}-2\psi_k)$$

$$A(f_k:p_k:s_{k+1}) = \sin(\phi_{k+1}-\phi_k) + Q_k \sin(\phi_k+\phi_{k+1}-2\psi_k)$$

$$A(s_k:p_k:f_{k+1}) = -\sin(\phi_{k+1}-\phi_k) + Q_k \sin(\phi_k+\phi_{k+1}-2\psi_k)$$

$$A(s_k:p_k:s_{k+1}) = \cos(\phi_{k+1}-\phi_k) - Q_k \cos(\phi_k+\phi_{k+1}-2\psi_k) \qquad (9)$$

and as before:

$$A(p:f_1) = \cos(\phi_1-\phi_o)$$
$$A(p:s_1) = \sin(\phi_1-\phi_o)$$
$$A(f_n:p) = \cos(\phi_p-\phi_n)$$
$$A(s_n:p) = -\sin(\phi_p-\phi_n)\,. \qquad (10)$$

Equations (9) and (10) represent the amplitude factors corresponding to each transition. To find the amplitude of any particular path the corresponding factors must be multiplied together.

While it is possible to derive equations (9) and (10) from first principles, similar to the derivation of the Solc filter rules, they are also directly accessible from the Solc results themselves. A plate-partial polarizer-plate combination is essentially the same as a three-plate Solc configuration. The middle plate, however, has zero retardation but introduces the additional attenuations, ρ_x and ρ_y, for paths along its respective axes. Since no time separation is introduced by the middle plate, the two paths which would normally be present are easily combined, resulting in the four path amplitudes, (9), instead of the eight which might otherwise be expected.

Equations (9) simplify considerably for some special cases. For example, in the perpendicular Lyot configuration discussed earlier the Q_k angles alternate between $+45^\circ$ and -45°. In that case, if the partial polarizers are all oriented at 0°, the terms of equations (9) simplify such that each factor has only one term:

$$A(f:p:f) = Q_k$$
$$A(f:p:s) = 1$$
$$A(s:p:f) = -1$$
$$A(s:p:s) = -Q_k\,. \qquad (11)$$

The particular case for which $\psi_k = 0$ for all polarizers in the system is, actually, a general configuration. In this case, the transmission is

a maximum since, when the plate retardations are multiples of the wavelength the plates are effectively absent. The transmission is only a function of the polarizers for those wavelengths, and is maximized when all polarizers are parallel.

In general partial polarizer designs are important, both theoretically and practically. The extra degrees of freedom may allow for some easily implementable realizations of useful transmission profiles. Additionally, relaxation of the perfect polarizer requirement, $\rho_y = 0$, usually results in a significant increase of ρ_x. This produces dramatic increases in the overall transmission efficiency of the filters.

At the present time a general synthesis procedure is being developed for partial polarizing filters. The extra degrees of freedom add some complexity over the Solc filter synthesis since they must be constrained either by restrictive assumptions or by optimization criteria. Both approaches are being used, and a systematic investigation into partial polarizing configurations will soon be possible.

(3) Transmission Profile Management

In the previous chapter a theoretical foundation for the three filter types has been presented. While some of the design criteria were mentioned, their implications and control were largely ignored. Some of these criteria, specifically bandwidth, free spectral range, secondary maxima, and overall transmission, will be discussed in this chapter. Other performance objectives such as the field of view and tunability will be deferred to subsequent chapters.

Of all the performance criteria, the easiest to control is the free spectral range (FSR). In the frequency domain the response is a periodic function, which is directly related to the impulse train response in the time domain. The periodicity in the frequency domain is, therefore, the reciprocal of the pulse position spacing in the time domain. Usually, the pulse position spacing will be the minimum distance between pulses in the impulse response. In general, however, the pulse position spacing will be the greatest common divisor of the interpulse spacings. If this greatest common divisor is Δt, then the corresponding free spectral range, FSR, will be

$$FSR = \frac{\lambda^2_o}{c\Delta t} . \tag{1}$$

Since there is a direct linear relationship between pulse spacings and crystal lengths, the time delay, Δt, will correspond to the greatest common divisor, d_o, of all the crystal lengths. Specifically

$$\Delta t = \frac{|n_o - n_e| d_o}{c} . \tag{2}$$

For a standard filter configuration all the crystal lengths are multiples of the shortest length. Most frequently then, the free spectral range is simply determined by the length, d_o, of the shortest crystal according to (1) and (2).

In order to increase the free spectral range of a filter it is necessary to make the shortest crystal even shorter. For a Lyot or partial polarizing configuration this usually means adding one more short element to the filter.

For a Solc, no internal polarizer, filter with identical plates, each crystal must be made shorter and the entire design reconfigured. Composite designs, in which a thin Lyot element is used as a blocking filter to increase the free spectral range of a Solc filter, are also feasible.

The design criterion which is usually of greatest concern is the bandwidth, or full width at half maximum (FWHM). Although bandwidth is intimately linked to secondary maxima, or sidelobe, level, it may be approximated separately. While the free spectral range is determined by the minimum interpulse spacing, the bandwidth depends mainly on the total time duration of the impulse response.

The minimum bandwidth response for a given total time duration, T, results from a single Lyot element of length

$$d_T = \frac{cT}{|n_o - n_e|} . \tag{3}$$

For this element the transmission is

$$\tau(\lambda) = \cos^2\left(\frac{\pi cT}{\lambda}\right) \tag{4}$$

and the full width at half maximum is

$$\text{FWHM} = \frac{\lambda_o^2}{2cT} \tag{5}$$

where λ_o is the operating wavelength.

As the crystal of length, d_T, is divided into shorter crystals to make a Lyot, Solc, or partial polarizing filter, the FWHM will increase somewhat while the secondary maxima recede. If the total length, d_T, is used to make an n element Lyot filter, the bandwidth becomes approximately:

$$\text{FWHM}_{\text{Lyot}} = 1.77 \frac{\lambda_o^2}{2cT} . \tag{6}$$

Similarly, if the total length is sliced into n equal length elements and used for a Solc fan or folded filter, the bandwidth becomes:

$$FWHM_{Solc} = 1.59 \frac{\lambda_o^2}{2cT} . \tag{7}$$

In general, except for extreme impulse response apodizations, the FWHM will be inversely related to the total crystal length and will be of the above form:

$$FWHM = K \frac{\lambda_o^2}{2cT} \tag{8}$$

where K is a constant, depending on the design. The constant, K, is always greater than one, and for most designs will be less than three.

The finesse, $\mathscr{F}$, of a filter is the ratio of its free spectral range to its bandwidth. Since the FSR is determined by the greatest common divisor of its crystal lengths and the FWHM is determined by the total crystal length, the finesse is a function of their ratio. Specifically,

$$\mathscr{F} = \frac{2T}{K\Delta t} . \tag{9}$$

For the Lyot and Solc filters,

$$\begin{aligned} \mathscr{F}_{Lyot} &\simeq 1.13 \cdot 2^n \\ \mathscr{F}_{Solc} &\simeq 1.26 \cdot n \end{aligned} \tag{10}$$

where n is the number of elements in each filter.

To be comparable to a Fabry-Perot filter whose finesse is usually between 20 and 50, a Lyot filter would have 4 to 6 elements and a Solc filter would have 20 to 50 plates. Other, more highly apodized designs, would have correspondingly more plates for a given finesse. Partial polarizing filters, for example, have about 25% more elements than a corresponding Lyot filter.

Each of the birefringent filter types has the freedom to control transmission profile parameters such as the bandwidth and sidelobe level. Any change in the filter design which introduces a taper in the impulse response will have a corresponding effect on the transmission profile shape. There are, therefore, two major questions. The first is, what are the relation-

ships between desirable transmission profiles and impulse response tapers. The second is, what impulse response tapers are available with birefringent filters and how are they achieved.

Although the first issue, the relationship of desirable spectral properties to impulse response, is an extremely sophisticated problem, it has received considerable attention recently. The impulse response characteristics of birefringent filters are identical to those of finite impulse response digital filters and discrete array antenna systems. Consequently the results from these disciplines are directly applicable to the design of birefringent filters. An extremely complete catalog of known filter responses is given by Harris (1978), though only selected designs and design criteria will be discussed here.

The fundamental design tradeoff is between bandwidth and sidelobe (secondary maxima) level. In general filter designs which result in impulse responses which peak in the center and taper downward towards the edges result in lower sidelobes but wider bandwidths. This tends to concentrate most of the filter transmission in the main passband, but at the expense of some resolution. Two similar, but slightly different design objectives are, first, to minimize the peak sidelobe level for a given bandwidth, and, second, to minimize the integrated out of band transmission for a given bandwidth. The first criterion is natural when the filter is to be used for examining the line structure of emission spectra in the presence of very low background radiation. For examining the structure of absorption spectra in the presence of a large background continuum, the second criterion is more applicable. It has been demonstrated that the first objective is achieved by Dolph-Chebyshev tapers (Harris, 1978) while the second is achieved by discrete prolate spheroidal sequences (Slepian, 1978). In addition, many useful suboptimal designs are known which have other desirable properties.

The question of which impulse response tapers are desirable may be considered to be solved for these purposes. This leaves the remaining question of which tapers are achievable by birefringent filters. In principle the answer to this question is that all tapers are achievable by birefringent filters. As discussed previously, Harris (1964) has presented a synthesis algorithm for achieving any impulse response with a Solc filter.

In particular, Schiffman and Young (1968) have demonstrated that this technique can, in fact, be used to build Dolph-Chebyshev birefringent filters. Figure 1 shows the response of these designs for 15-plate filters with sidelobe levels of 10^{-1}, 10^{-2}, 10^{-3} and 10^{-4}. An interesting characteristic of these profiles is that, while the peak sidelobe level is minimized, all the sidelobes are equal.

A corresponding synthesis for filters with perfect internal polarizers was presented by Ammann and Chang (1965). It demonstrates that any specified impulse response may be achieved by such a filter. The disadvantage of this procedure is that the internal polarizers decrease the transmission substantially from that achievable with a lossless Solc filter. As an advantage, however, is the possibility of using fewer crystals, of differing lengths, such as in the Lyot design. The number of polarizers is not always prohibitive, therefore.

Practical considerations impose some restrictions on the list of achievable responses. For Solc type, lossless filters, the large number of elements required is the major problem. With larger numbers of plates, the angle tolerances for setting the plates become more severe. Present technology for producing precisely matched elements, coupled with the computational capability for initial alignment, makes these filters more practical now than in the past. Designs with more than 50 plates, however, are probably only speculative at this time.

Perfect internal polarizing filters are limited by the number of required polarizers, and by the difficulties of manufacturing elements of different but precisely matched lengths. As will be discussed subsequently, tunability allows for the compensation of length mismatches, so that problem may be avoided. The transmission loss of polarizers, however, presents a limitation in the use of these filters for arbitrary spectral transmission profiles. The solution to the practical limitations has been to use feasible design techniques which result in suboptimal, but satisfactory, profiles. A very common design modification for Lyot filters is the introduction of a contrast element, an element whose length equals the second longest element in the original Lyot filter. This introduces an amplitude taper which reduces the sidelobe level significantly. In Figure 2 a four-

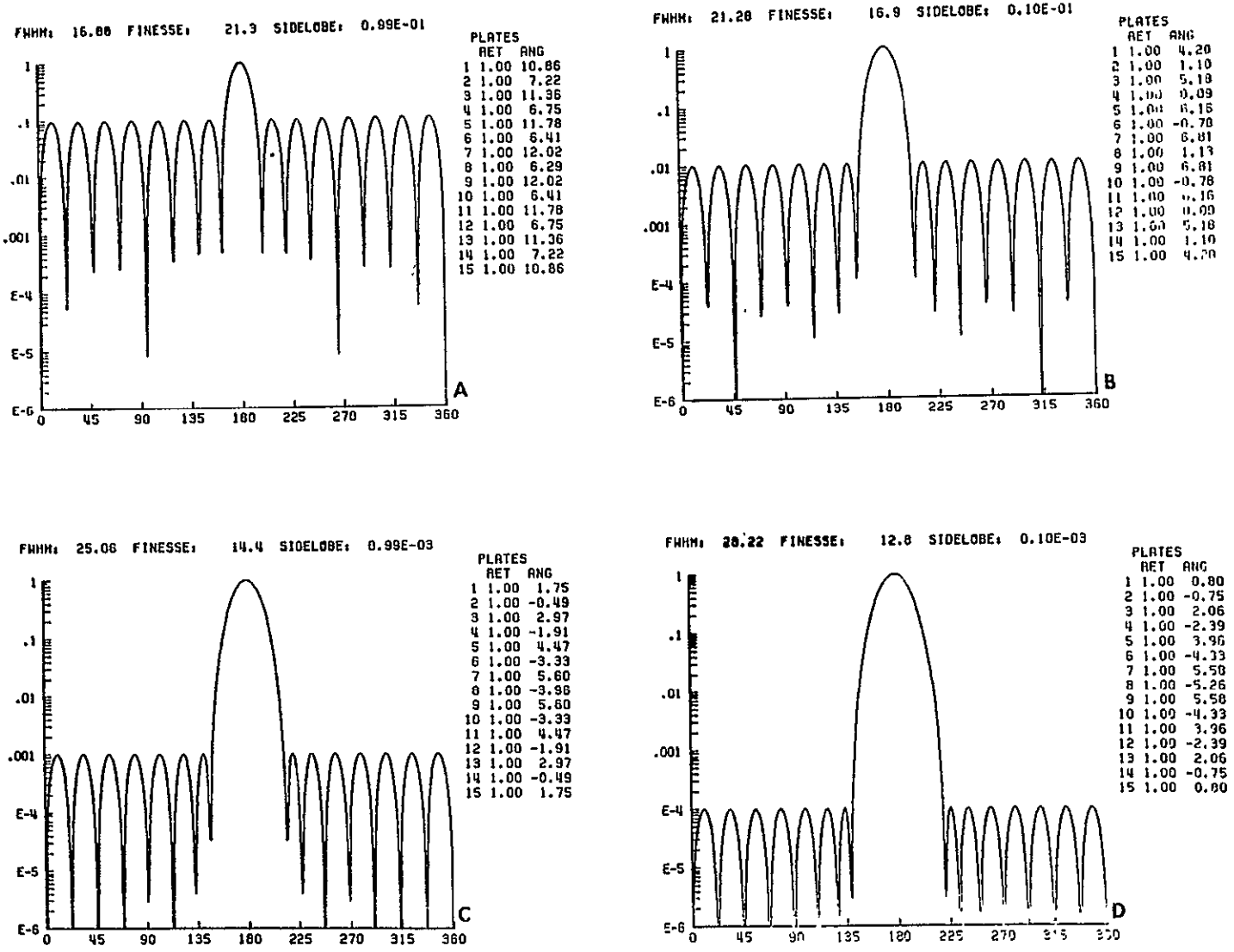


Figure 1. Transmission versus frequency for 15-plate Dolph-Chebyshev profiles with sidelobe levels of 10^{-1}, 10^{-2}, 10^{-3}, and 10^{-4} (a,b,c,d).

element Lyot response is compared with a five-element filter which uses a contrast element. Also shown in Figure 2 is the response of an untapered filter whose total length is equal to the contrast element filter.

Partial polarizing filters, having more degrees of freedom, certainly have all the capabilities of both perfect polarizer and no polarizer filters. Their utility is that because of the extra degrees of freedom, more practical realizations for filters may be found. In particular, designs which use Lyot element lengths or nearly Lyot lengths but which have better sidelobe control have been the major interest. The Lyot element lengths are the most efficient since they represent the division of a total crystal length into the fewest pieces while achieving a given minimum length, $c\Delta t$. Any division using fewer plates will have zeros in the interior of its impulse response.

The simplest partial polarizing design is one in which the first internal polarizer of a Lyot filter is replaced by a partial polarizer with:

$$\rho = \rho_y/\rho_x = 1/3 . \tag{11}$$

The response of such a design is very similar to that of a contrast element Lyot filter. A four-element Lyot filter with one partial polarizer is shown in Figure 3. The response of that filter should be compared with the contrast element Lyot shown in Figure 2. There is a 26% increase in bandwidth but this is due to the 27% more crystal in the contrast element filter. The sidelobe levels are comparable, however.

If all the polarizers are replaced by partial polarizers in a Lyot filter the profile does not improve. There are, however, designs in which more than one polarizer can be made partial and improvement is expected. These are discussed by Title (1976) as the alternate partial polarizing (APP), double partial polarizing (DPP), and redundant alternate partial polarizing (RAPP), filters.

The APP design is one in which alternate internal polarizers are partial, $\rho = .333$, and the others are perfect, $\rho = 0$. Across a partial polarizer the crystal lengths are in the Lyot ratio, 2:1, but across a perfect polarizer the ratio 3:2 is needed. A four-plate alternate partial polarizer filter is

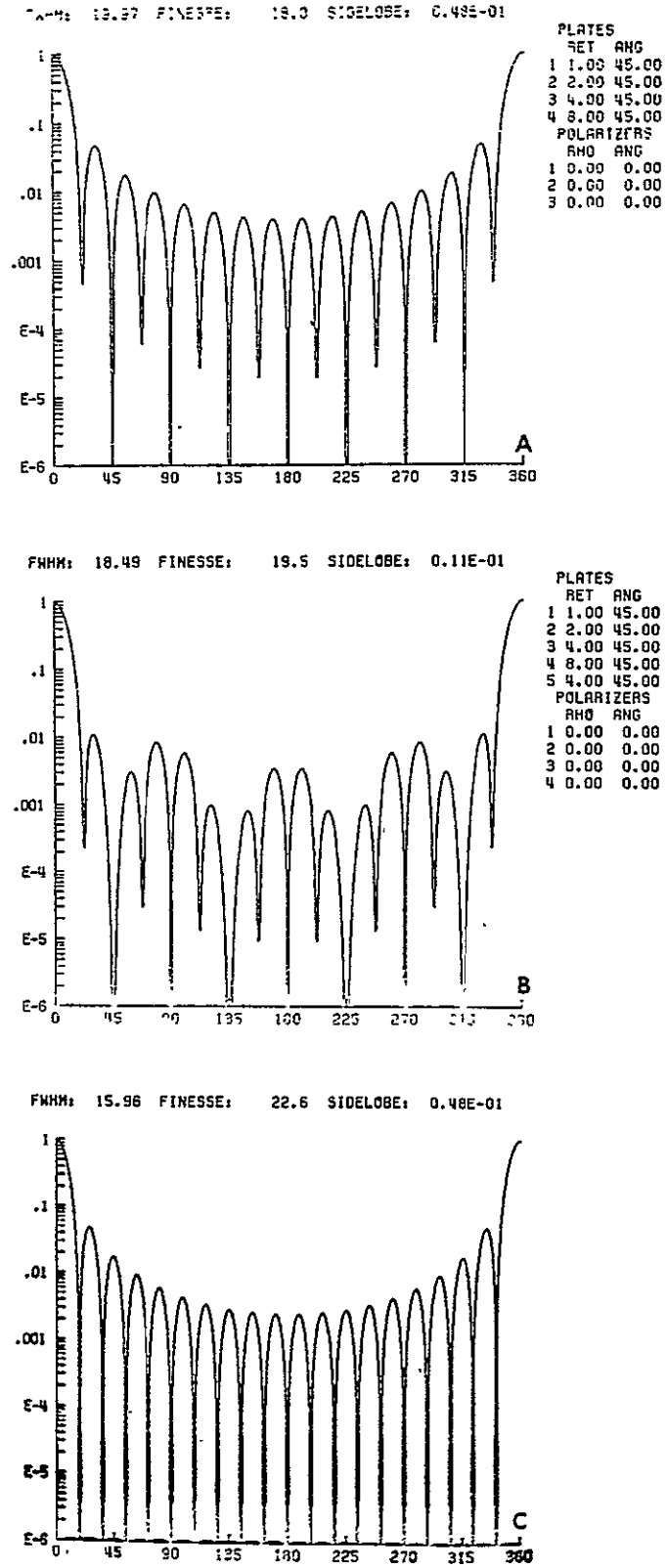


Figure 2. Transmission versus frequency for four-element Lyot (a), contrast element Lyot (b), and linear pulse response filter with the same length as the contrast Lyot (c).

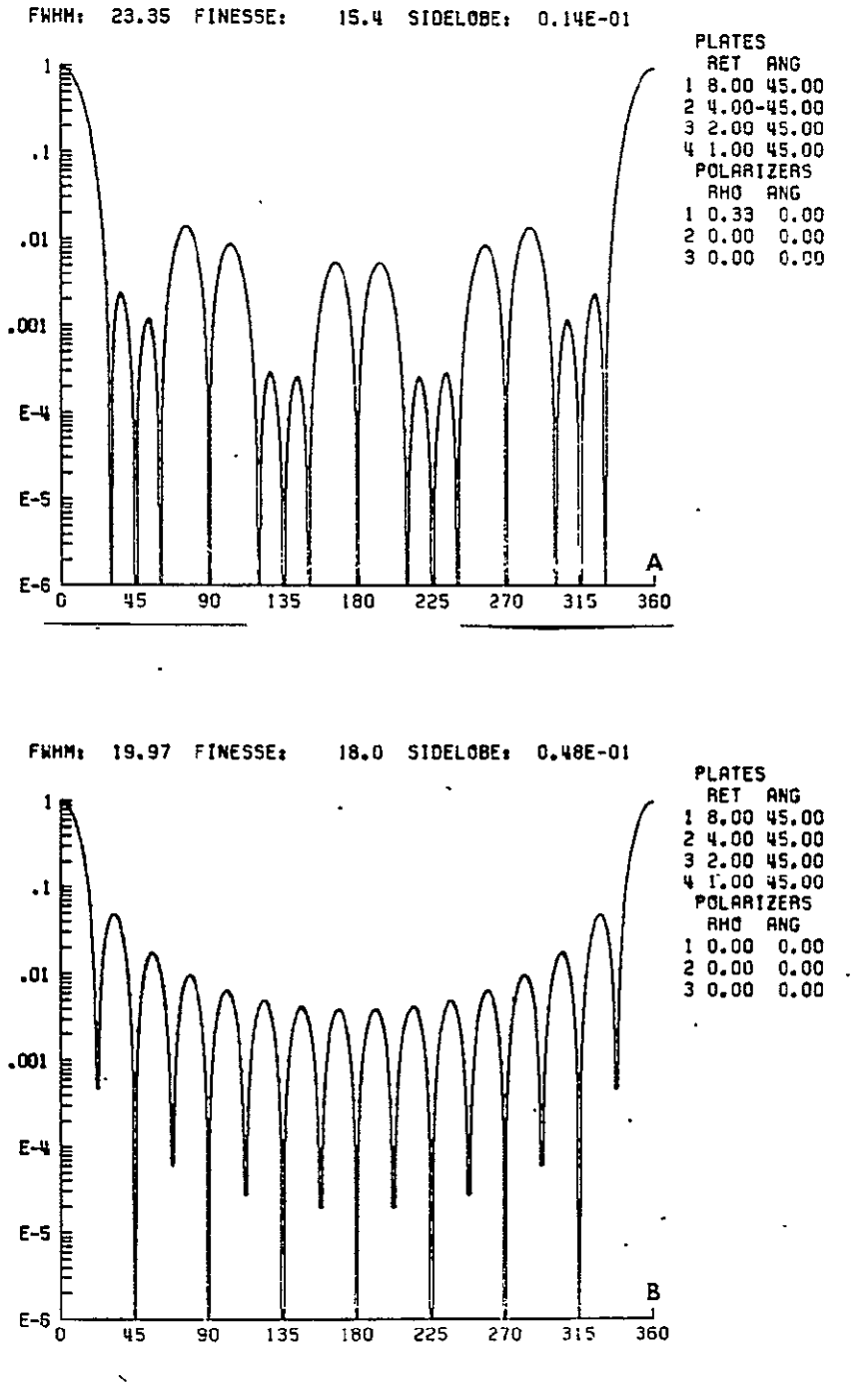
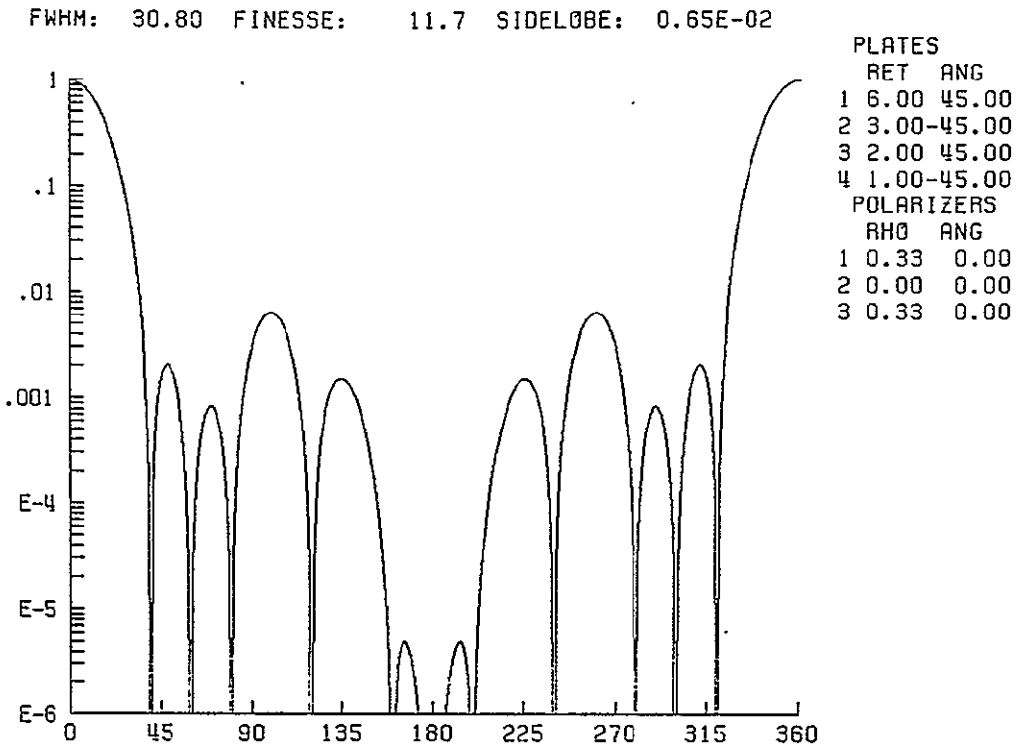


Figure 3. Transmission versus frequency for a four-element Lyot with a single partial polarizer (a) and a pure Lyot (b).

shown in Figure 4. As is typical of partial polarizer designs, the plates following the partial polarizers are orthogonal to the plates preceding the partial polarizers. Notice the extreme sidelobe reduction, while the bandwidth expansion is due to the reduced total crystal length.

A double partial polarizing module may be designed with three Lyot length elements arranged in the order (2:4:1) with partial polarizers between elements. For this arrangement the partial polarizers should have orthogonal transmissions, ρ, equal to .333 and .111 respectively. Figure 5 shows the responses of a single DPP module and, for comparison, the equivalent perfect polarizer Lyot filter. Notice that the sidelobe level has been reduced but at the expense of an increase in bandwidth. A filter can be constructed of double partial polarizing modules alternating with perfect polarizers just as an APP filter is constructed from single partial polarizing modules.

A partial polarizing design which uses the idea of duplicating element lengths, like the contrast element design, is the redundant alternate partial polarizing configuration. The elements have Lyot lengths, but all but the longest and shortest elements are duplicated. The polarizers alternate, partial and perfect. Figure 6 shows a RAPP configuration with six elements, and the same configuration with perfect polarizers. Notice the excellent sidelobe reduction at only a modest bandwidth increase. In addition to their spectral profile advantages, the partial polarizing designs have been found to be much less sensitive to construction errors than the traditional Lyot, perfect polarizer, design. In particular, crystal length errors produce much weaker secondary maxima in the partial polarizing designs than in the Lyot filter. Also, surprisingly, some partial polarizing designs have been found to be particularly tolerant to the polarization, ρ. This is particularly important for short wavelengths where sheet polarizing materials have poor polarizing properties. Specifically the polarization of sheet polarizers deteriorates rapidly at wavelengths below 4500 Å so a design which is robust with respect to ρ is very desirable. Figure 7 shows the behavior of a partial polarizer design whose elements have lengths (4:8:6:4:2:1). Notice that as ρ increases the finesse decreases, and the sidelobe structure changes, but the filter does not deteriorate catastrophically. A current topic is the search for partial polarizing designs which have their best performances when ρ is about .25. That should insure a greater uniformity over a broad range of ρ. Designs of this type should make it possible to design filters which are



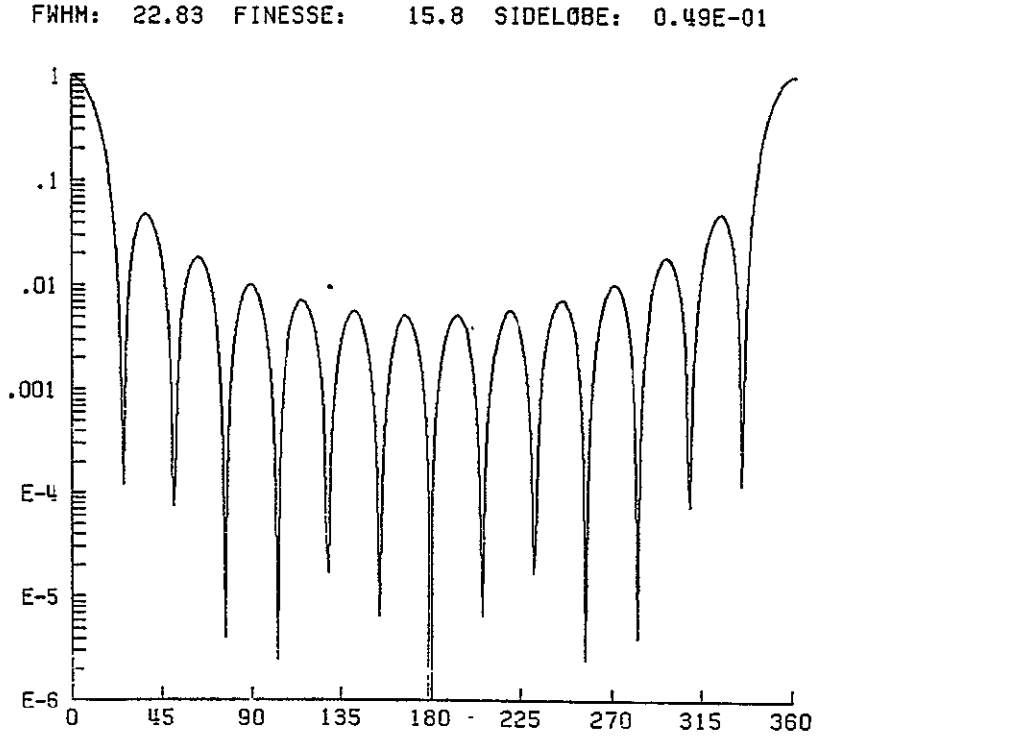


Figure 4. Transmission versus frequency for a four-element APP filter (a) and a uniform pulse filter of the same length (b).

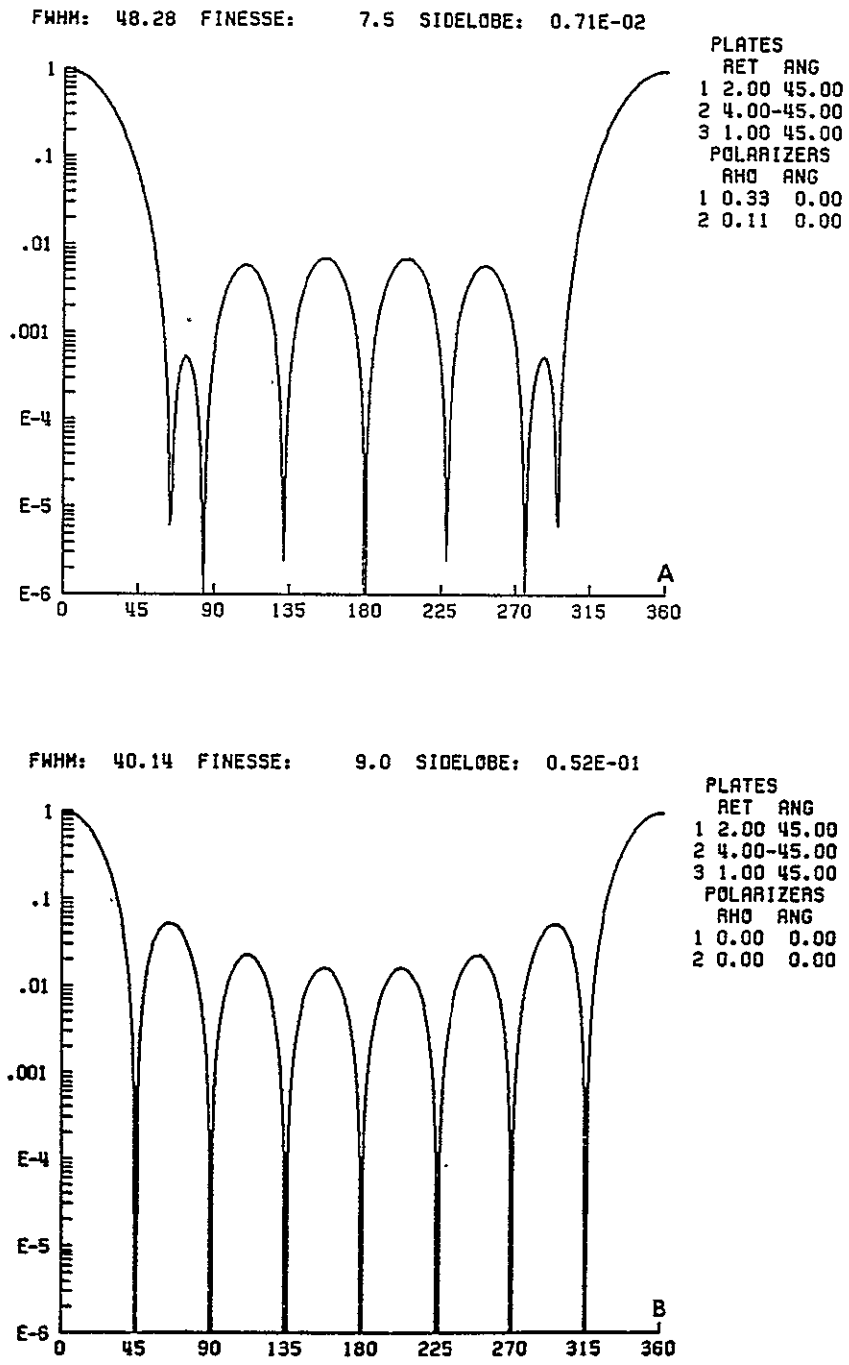


Figure 5. Transmission versus frequency for DPP (a) and three-element Lyot (b).

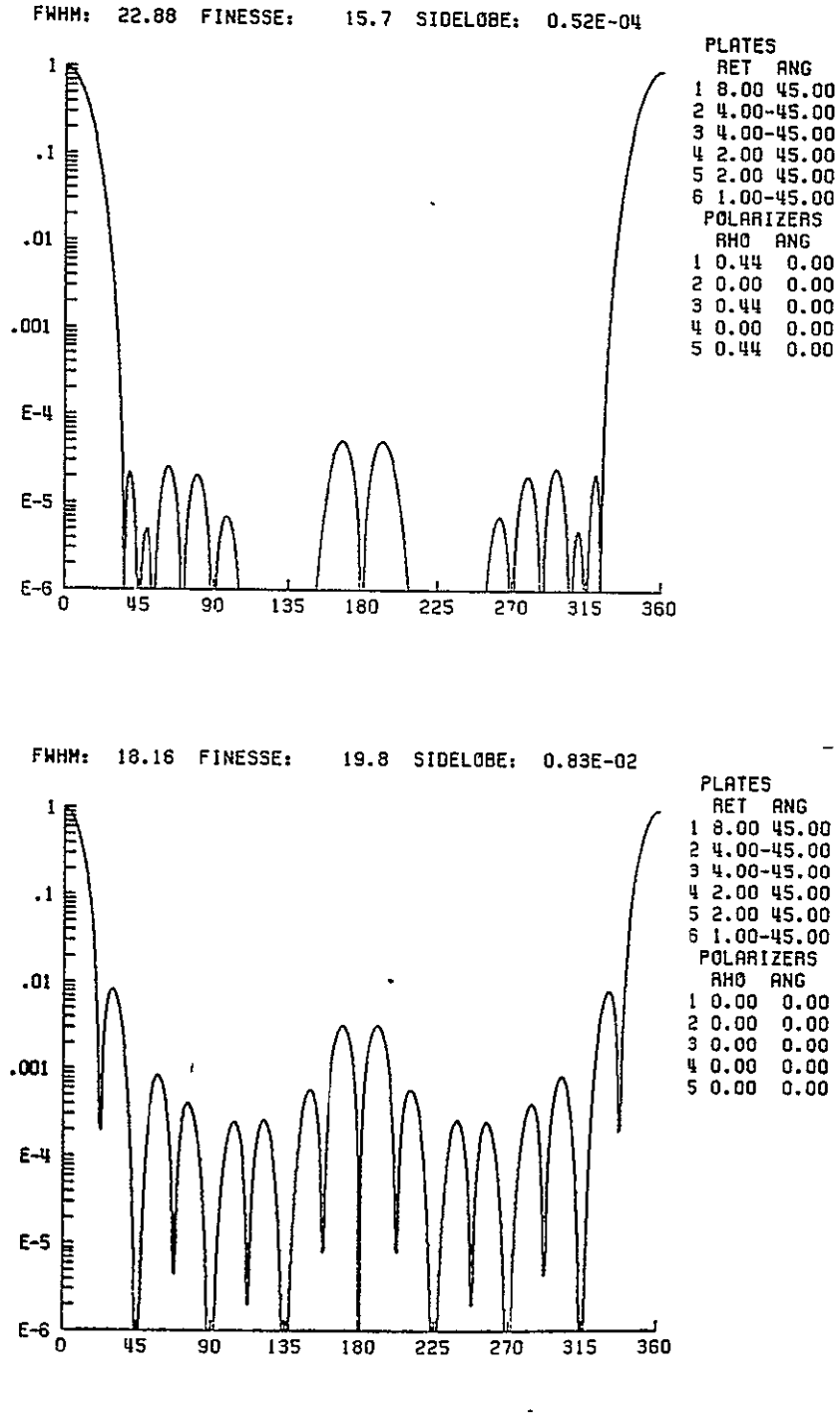


Figure 6. Transmission versus frequency for a six-element RAPP (a) and the same crystal configuration with perfect polarizers (b).

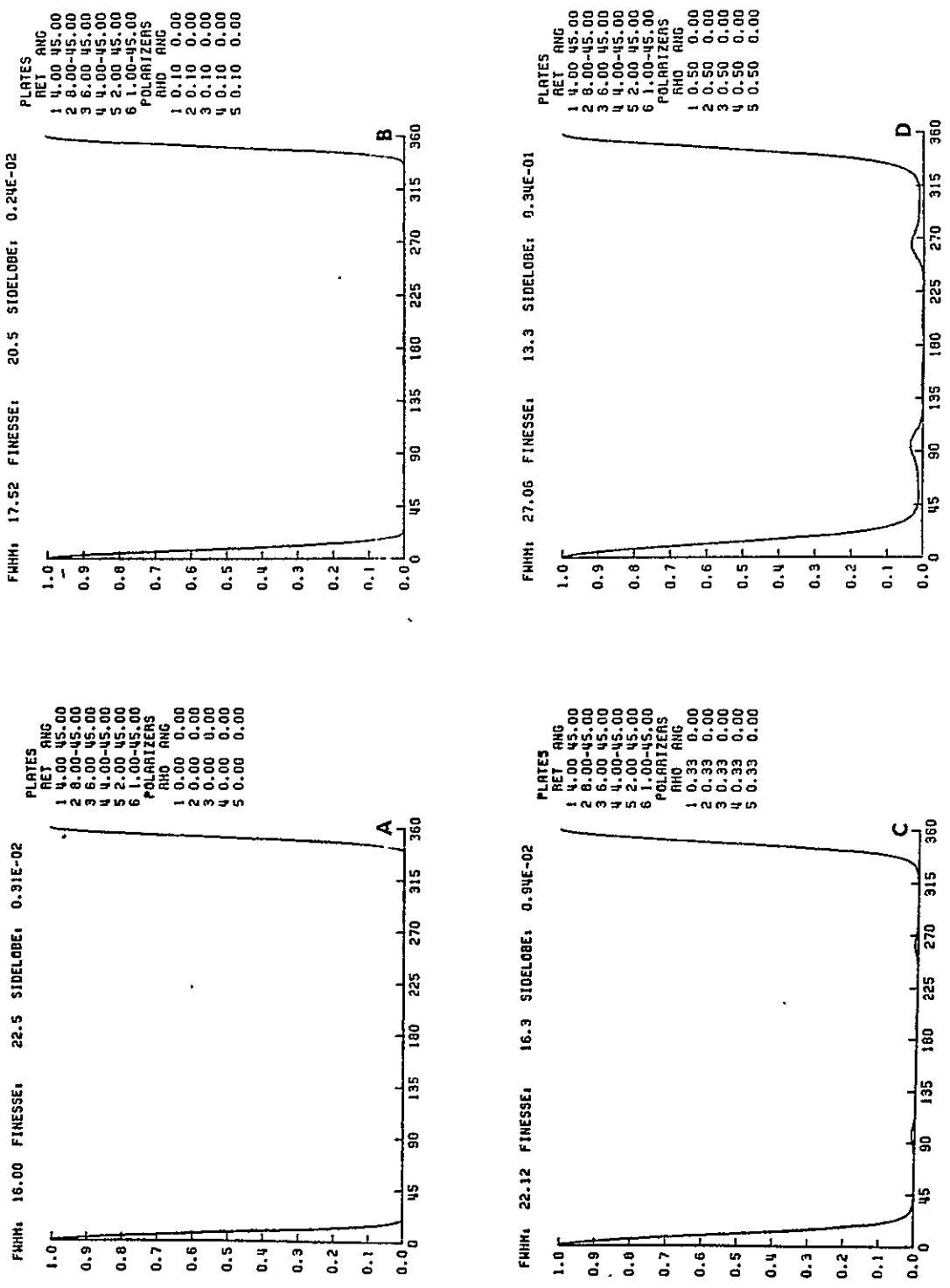


Figure 7. Transmission versus frequency for (4:8:6:4:2:1) plate ratios with ρ=0., .1, .33 and .5 (a,b,c,d).

acceptable down to 3500 Å using the same materials and construction techniques as used for longer wavelengths.

The major transmission losses in birefringent filters are caused by absorption by polarizers and imperfect filter construction. A minor source of loss is interface Fresnel reflection at the plates. However, for filters that are in an index matching fluid this problem is minimal. A ten percent transmission loss occurs for 100 calcite or 3400 quartz surfaces.

All birefringent filters require high quality entrance and exit polarizers. These can be ordinary sheet polarizer, but it is also practical to use thin film or crystal polarizers which can have virtually 100% efficiency. Unfortunately both thin film and crystal polarizers are nearly cubical and a serious length penalty is incurred if such polarizers are used internally in the filter. The additional length introduces optical complications, so that for most designs sheet polarizing materials must be considered the only option for the majority of the internal polarizers.

A perfect polarizer filter requires (n-1) internal polarizers for elements, so that relative to a lossless filter the transmission is

$$\tau_{Lyot} = \tau_o^{n-1} \tag{12}$$

where τ_o is the transmission of the internal polarizers in the pass direction. For a partial polarizing filter

$$\tau_{Partial} = \prod_{i=1}^{n} \tau_i$$

where the τ_i are the transmissions of the partial polarizers. Figure 8 is a plot of total transmission versus number of internal polarizers for various τ_o. This should be compared with Figure 9 which shows the finesse of a Lyot filter as a function of the number of internal polarizers. Figure 10 has plots of transmission, τ_o, versus wavelength for high quality specially treated polarizer materials mounted in AR coated glass. Visible wavelength polarizers superior to those in Figure 10 can be produced, but at present are not stable for more than a few months. Beyond 7500 Å there are at least two methods to produce good polarizers besides polarizing sheet.

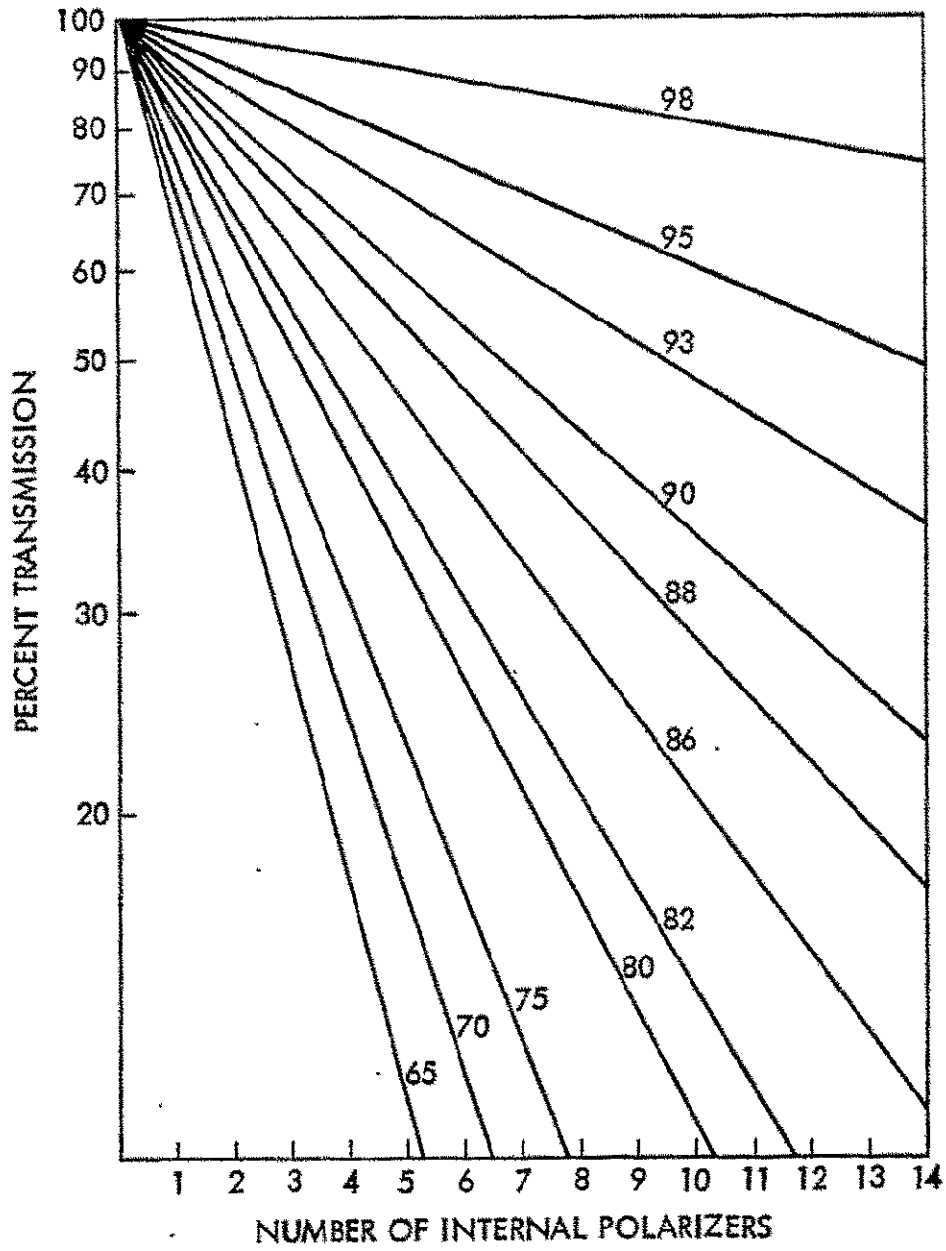


Figure 8. Total transmission of a Lyot filter versus number of internal polarizers.

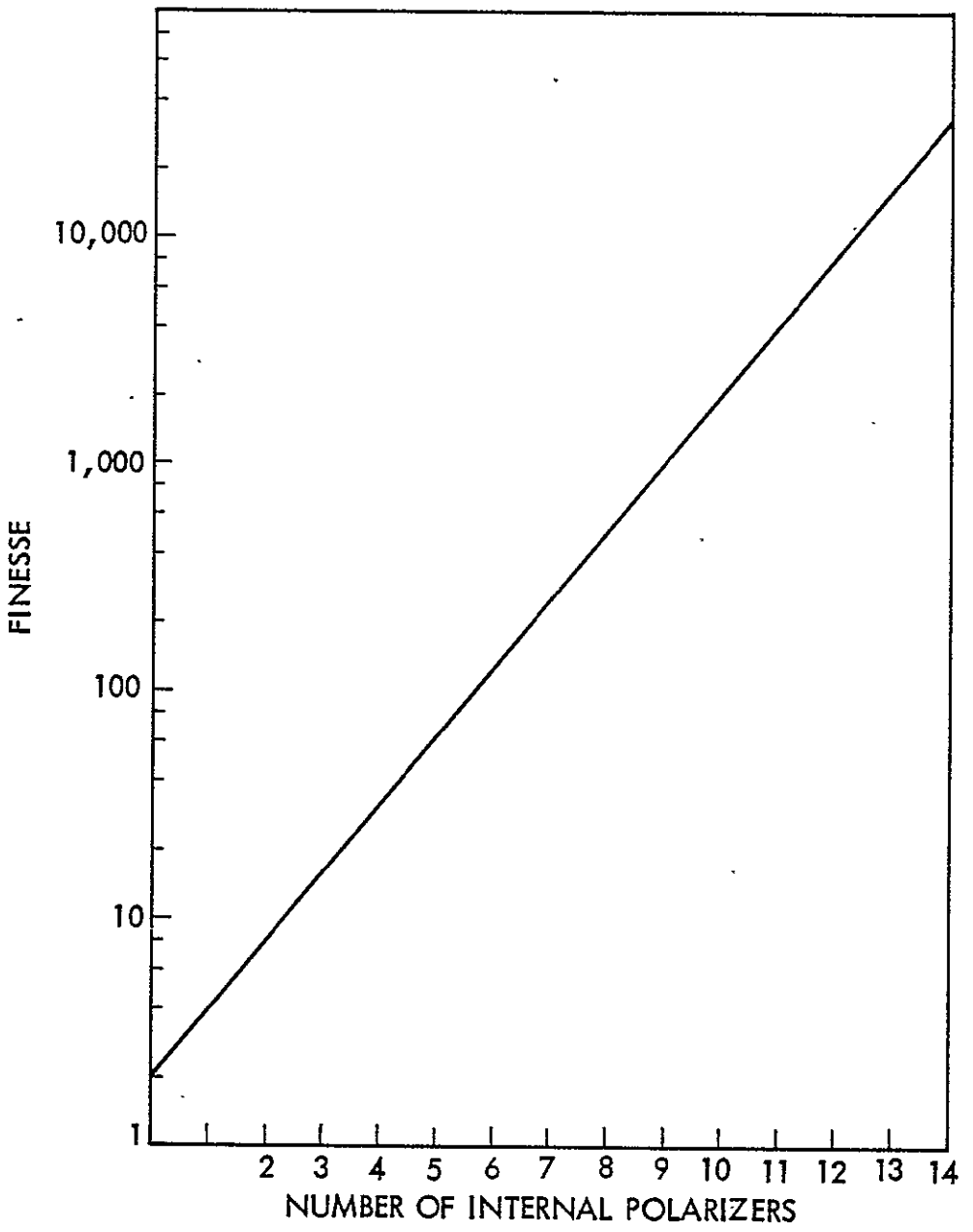


Figure 9. Finesse of a Lyot filter versus number of internal polarizers.

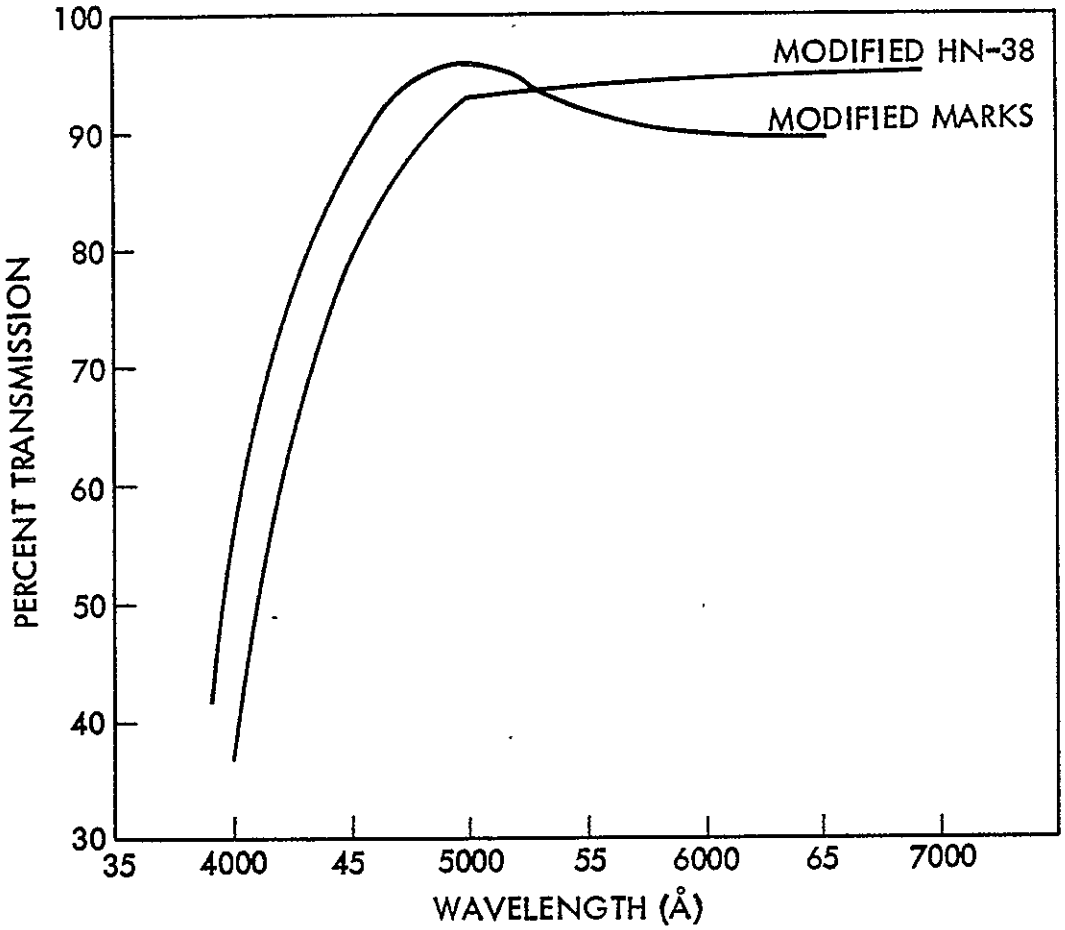


Figure 10. Transmission versus wavelength of specially treated polarizing materials.

These are sheet crystal polarizers and wire grids. At present it would appear possible, using modern holographic techniques, to make wire grid polarizers useful for wavelengths beyond one micron. Below 4000 Å, however, there is a serious problem with polarizers. UV polarizing sheet does exist, but its maximum transmission is in the region of 60%. It might be possible to create partial polarizer sheet material below 4000 Å with reasonable transmission, which would make the partial polarizer designs very attractive. Another possibility is the use of electron beam and X-ray lithographic techniques for production of wire grid polarizers. At the present state of the art it might be possible to create polarizers for wavelengths down to 1500 Å.

For the visible Figures 8 through 10 provide a useful guide to what can be done using Lyot type designs. If a total transmission criterion is at least ten percent transmission, then operation around 4000 Å is limited to a finesse of 64. At present the best blocking filters have FWHM of 6 Å in this spectral region, so that a minimum FWHM of the birefringent filter is 0.1 Å. On the other hand it is possible to construct a filter with a finesse of 2048 which operates from 4600 Å to 7500 Å with a minimum transmission of 35% and a maximum transmission of 64%. If the wavelength range is restricted to above 5000 Å, a transmission of 48 to 60 percent can be achieved.

At present it is reasonable to construct 0.1 Å filters for the visible and difficult to produce filters narrower than .05 Å. A 0.1 Å filter could be blocked for the entire range 4800 - 7800 Å and have a transmission in the range of 25% to 50% over the entire spectral band. For a more moderate resolution of say 2 Å, the required finesse implies a transmission range of 35% to 65%. And for a 25 Å filter, the transmission range would be 53% to 75%.

Polarizer limitations make it necessary to consider other designs which try to combine the transmission efficiency of Solc filters with the crystal utilization of Lyot filters. As has been demonstrated in the previous section it is often possible to find Solc type filters with some multiple length plates. A particularly dramatic example of this is shown in Figure 11. The nine-plate Solc filter with a triangular taper impulse response is shown, along with a five-plate compression. Notice that the first and last three plates have been replaced by single, length three, plates at their average orientations. The performances of these two filters are equivalent, but the compressed filter has only half as many plates. In Figure 11(c) this five-element

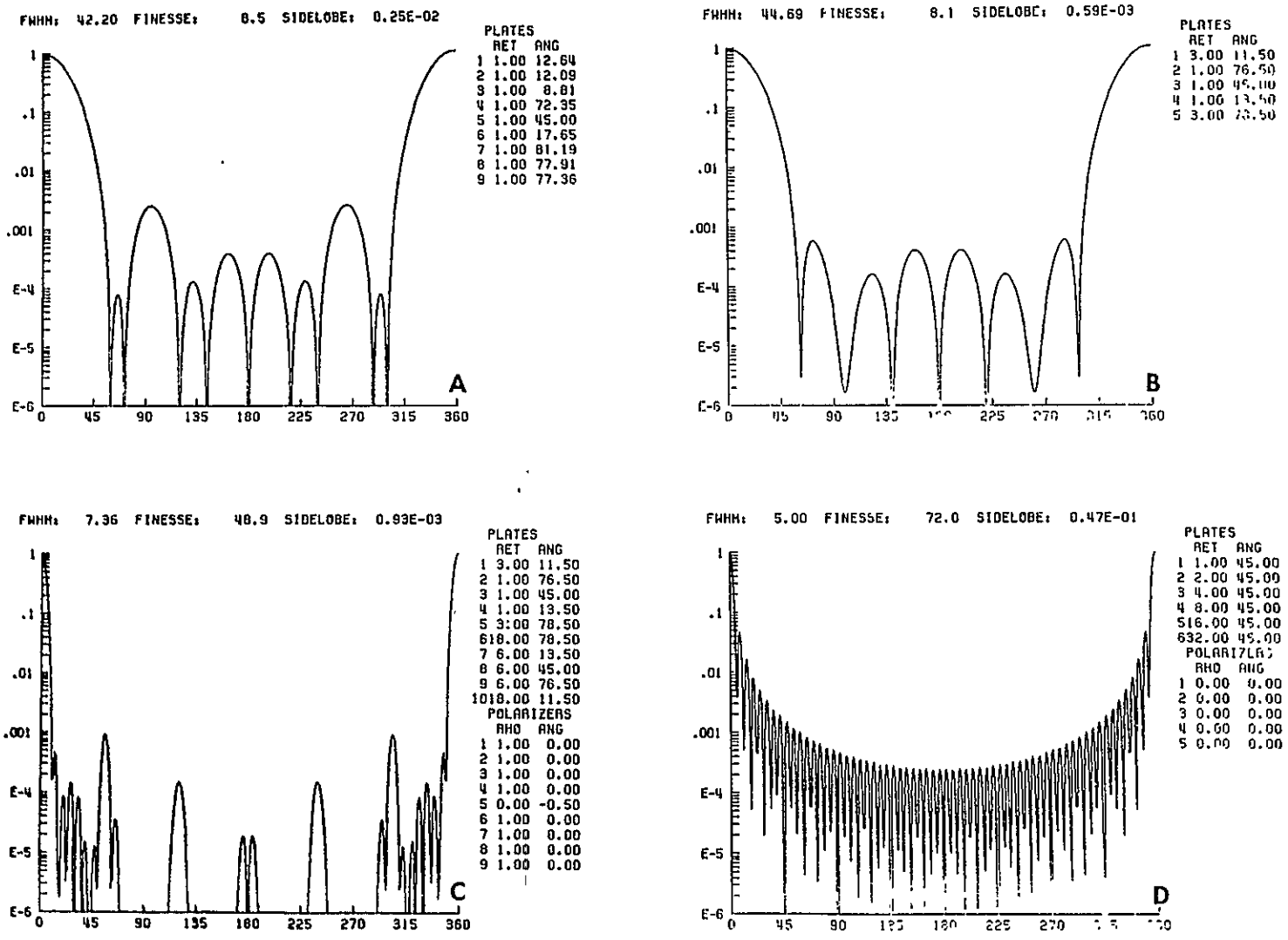


Figure 11. Transmission versus frequency for nine-plate liner taper Solc (a), five-plate approximation (b), two five-element groups with element ratios of 6:1 (c), and a six-element Lyot (d).

filter has been used as a module and duplicated using plates six times as long. The two sections are separated by a perfect polarizer which combines the large free spectral range of the short section with the narrow bandwidth of the longer section. The resulting filter compares favorably with the six-element Lyot filter which has the same total crystal length. It has ten elements instead of six, a finesse of 49. instead of 72., but it only has one internal polarizer instead of five, and sidelobes which are a factor of 50 below the Lyot. This or similar filter designs would be very attractive for the ultraviolet or infrared wavelengths where polarizers are critical.

(4) Tuning Birefringent Filters

There are two methods for tuning birefringent filters, a direct method which is conceptually simple but difficult in practice and an indirect method which is practical but difficult to understand. The direct method is to actually change the optical length of each crystal plate in order to make each plate have a retardation which is a multiple of the wavelength of interest, while maintaining the proper length ratios. This may be accomplished by using pairs of wedge shaped plates as adjusters for each element. By moving the wedges in and out, the thickness in the optical path may be changed and the elements may be tuned. The optical path length may also be changed by controlled temperature changes. A change in temperature will change both the physical length and the birefringence. Although direct adjustment is the method for tuning air spaced Fabry-Perot and Michelson interferometers it has not been the most practical method for birefringent filters.

In practice most birefringent filters are tuned by rotation of thin waveplates or polarizers. The key to understanding this method of tuning is to consider the effect of a birefringent plate on the polarization state of monochromatic light. Thus spectral or frequency domain analysis is more suited to this problem than the pulse response or time domain techniques of previous chapters.

Consider linearly polarized light incident on a birefringent element as in a Lyot filter. The plane of polarization is assumed to be at 45^{o} to the fast axis of the plate and the incident light is monochromatic with wavelength, λ. In the coordinate system of the plate the incident light may be expressed as:

$$A_f^i = \frac{1}{\sqrt{2}} \cos\ (2\pi \frac{c}{\lambda} t)$$

$$A_s^i = \frac{1}{\sqrt{2}} \cos\ (2\pi \frac{c}{\lambda} t). \qquad (1)$$

The birefringent plate introduces a pure time delay, Δt, in the slow axis, which is equivalent to a wavelength dependent phase shift

$$\Delta\phi = 2\pi c \frac{\Delta t}{\lambda} . \tag{2}$$

The light exiting the plate is, therefore, represented as

$$A_f^e = \frac{1}{\sqrt{2}} \cos\ (2\pi \frac{c}{\lambda} t)$$

$$A_s^e = \frac{1}{\sqrt{2}} \cos\ (2\pi \frac{c}{\lambda} t - \Delta\phi) . \tag{3}$$

When the phase shift, $\Delta\phi$, is a multiple of 2π the two components of (3) are in phase and the output is linearly polarized parallel to the incident plane of polarization. If $\Delta\phi$ is an odd multiple of π, the output is again linearly polarized, but in a plane perpendicular to the input plane. For other values of $\Delta\phi$, the output light is, generally, elliptically polarized. The axes of the ellipse are parallel and perpendicular to the input polarization plane. Rotating the output coordinate system by 45^o to express the output in terms of components parallel and perpendicular to the entrance polarizer:

$$A_{||} = A_f + A_s$$

$$A_{\perp} = -A_f + A_s . \tag{4}$$

After some algebra (3) and (4) imply:

$$A_{||} = \cos \frac{\Delta\phi}{2} \cos\ (2\pi \frac{c}{\lambda} t - \frac{\Delta\phi}{2})$$

$$A_{\perp} = \sin \frac{\Delta\phi}{2} \sin\ (2\pi \frac{c}{\lambda} t - \frac{\Delta\phi}{2}) . \tag{5}$$

Equations (5) may be combined, eliminating the time dependence, to exhibit the ellipticity

$$\left(\frac{A_{||}}{\cos \frac{\Delta\phi}{2}}\right)^2 + \left(\frac{A_{\perp}}{\sin \frac{\Delta\phi}{2}}\right)^2 = 1 . \tag{6}$$

Equation (6) is, as asserted, the equation of an ellipse whose axes are parallel and perpendicular, respectively, to the entrance polarizer, but whose ellipticity is a function of the wavelength.

In order to tune this element, therefore, the plate must be followed by an optical element which is sensitive to the ellipticity of the output. A polarizer is a detector of linearly polarized light, and is therefore sensitive to the limiting cases:

$$\Delta\phi_{\parallel} = 2n\pi$$
$$\Delta\phi_{\perp} = (2n-1)\pi . \tag{7}$$

Using a polarizer and switching from the parallel to the perpendicular orientation switches the wavelength from

$$\lambda_{\parallel} = \frac{c\Delta t}{n}$$
$$\text{to} \qquad \lambda_{\perp} = \frac{2c\Delta t}{2n-1} \tag{8}$$

A more general ellipticity sensitive device is constructed by inserting an achromatic quarter waveplate between the element and the polarizer as shown in Figure 1. An achromatic quarter waveplate is a birefringent plate which behaves as implied by equations (3) but for which the phase shift, $\Delta\phi = \pi/2$, independent of wavelength. The theory of these waveplates is described in Title (1975) and McIntyre and Harris (1968) which are included in Volume II, and in Pancharatnam (1955a and 1955b) and will not be described in this section. It is, however, the ease of making achromatic waveplates inexpensively which makes tunable birefringent filters practical.

While the theory underlying the construction of achromatic quarter waveplates is difficult, their application as tuning elements is not too difficult. Consider linearly polarized light incident on a quarter waveplate at some arbitrary orientation, θ, with respect to the fast axis of the plate. In the coordinate system of the plate the incident light is represented as:

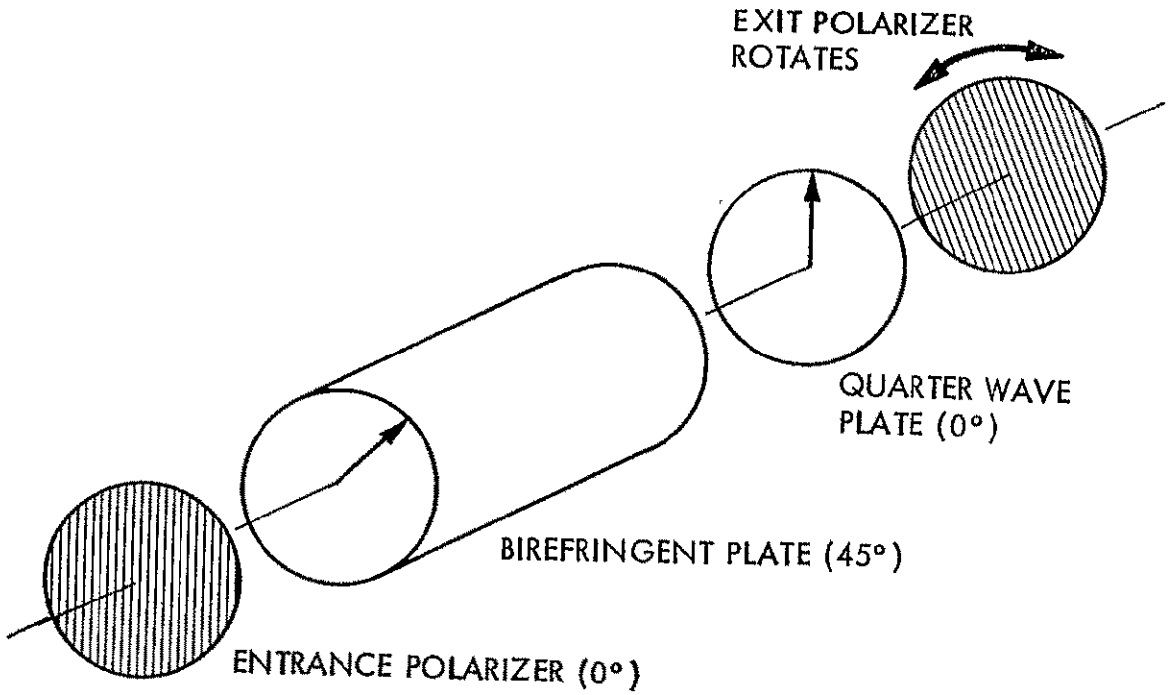


Figure 1. Diagram of simple tuning element.

$$A_f^i = \cos\Theta \cos\left(2\pi \frac{c}{\lambda} t\right)$$

$$A_s^i = \sin\Theta \cos\left(2\pi \frac{c}{\lambda} t\right) . \tag{9}$$

The light exiting the quarter waveplate is,

$$A_f^e = \cos\Theta \cos\left(2\pi \frac{c}{\lambda} t\right)$$

$$A_s^e = \sin\Theta \cos\left(2\pi \frac{c}{\lambda} t - \frac{\pi}{2}\right) \tag{10}$$

or $$A_s^e = \sin\Theta \sin\left(2\pi \frac{c}{\lambda} t\right) .$$

Equations (10) represent elliptically polarized light whose axes are the fast and slow axes of the waveplate. The ellipse so defined is represented:

$$\left(\frac{A_f^e}{\cos\Theta}\right)^2 + \left(\frac{A_s^e}{\sin\Theta}\right)^2 = 1 . \tag{11}$$

A quarter waveplate, therefore, converts linearly polarized light into elliptically polarized light. By reciprocity, however, it must also convert elliptically polarized light into linearly polarized light. The crucial point is that the plane of polarization of the linearly polarized light will depend only on the ellipticity as determined by (11).

The mechanism is, therefore, that each wavelength (modulo a free spectral range) will emerge from the birefringent element as elliptically polarized light with axes parallel and perpendicular to the input plane of polarization. An achromatic quarter waveplate with its fast axis parallel to the entrance polarizer, at 45° to the fast axis of the element, will convert the elliptically polarized light to linearly polarized light at an orientation which depends on wavelength. Specifically, from (2), (6), and (11), the dependence is:

$$\cos\Theta = \pm \cos\frac{\Delta\phi}{2} \tag{12}$$

or $$\Theta = \frac{\Delta\phi}{2} - n\pi$$

$$= \frac{\pi c \Delta t}{\lambda} - n\pi \ . \qquad (13)$$

At $\Theta = 0$ the natural wavelength, λ_o, of the element emerges, so (13) may be written

$$\Theta = \pi c \Delta t \left(\frac{1}{\lambda} - \frac{1}{\lambda_o}\right) \qquad (14)$$

The wavelength, λ, is then selected by rotating an exit polarizer to the appropriate angle, Θ.

The tuning scheme described above is the most commonly used system for Lyot filters. It has, however, one serious liability. Since the exit polarizer for the first element is the entrance polarizer for the second element, the second and subsequent birefringent plates must rotate as the filter is tuned. While rotating the crystals is theoretically satisfactory, as a practical matter it is very bad. No crystal elements are perfect, but the overall imperfections of a fixed set of crystals can be calibrated or polished out. If the crystals rotate, differentially, the errors become wavelength dependent and are very much more difficult to compensate. Additionally, if the rotation axes differ from the optic axis, displacements occur which are also wavelength dependent.

The problem of rotation of successive elements can be solved by the addition of an achromatic half waveplate in the tuning element. An achromatic half waveplate is a birefringent plate which behaves as implied by equations (3) but for which the phase shift, $\Delta\phi = \pi$, independent of wavelength. Conceptually they are just two quarter wave plates in series. When linearly polarized light is incident on a half waveplate at an orientation, Θ, with respect to its fast axis it exits linearly polarized but at an orientation, 2Θ. The half waveplate may, therefore, be used to rotate the plane of polarization in between the exit polarizer and the succeeding crystal. This configuration is shown in Figure 2.

The only problem with the position of the half waveplate in Figure 2 is that as the polarizer rotates to tune the element the half waveplate

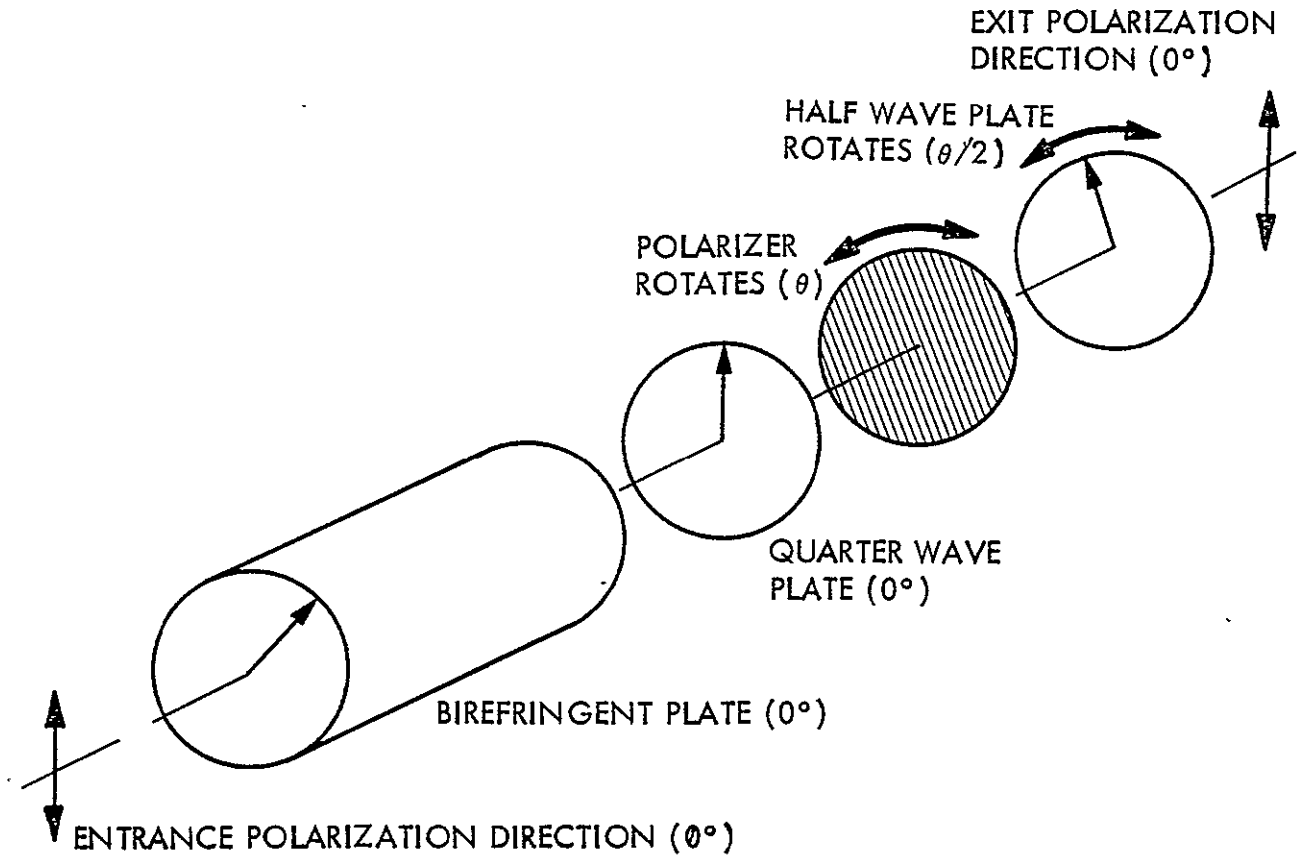


Figure 2. Diagram of a tuning element with a half wave plate behind it to rotate the exit polarization direction parallel to input direction.

must rotate half the angle in order to correct the orientation. If, however, the half waveplate is placed before the exit polarizer, as in Figure 3, only the half waveplate need rotate. The exit polarizer and the succeeding crystal will stay fixed. Each of these placements of the half waveplate have been used in the Lockheed tunable filters.

The tuning technique for partial polarizing and Solc filters is more complex than that for Lyot filters. In Lyot filters the tuning only has to be concerned with one plane of linear polarization since the entrance and exit polarizers reject the other plane. In the context of the configuration shown in Figure 3, the element is tuned to a wavelength, λ, if vertically polarized light at the entrance polarizer emerges vertically polarized at the exit polarizer. This is the behavior which was demonstrated above. The Solc and partial polarizing filters, however, place more demands on the tuning section. They require linearly polarized light of wavelength, λ, at any orientation, to emerge from the tuning section at the same orientation as the input. The polarizers are, of course, absent.

In order to determine the effect on arbitrary orientations it is only necessary to examine the effect on horizontal and vertical polarizations. All other polarizations are simply linear combinations of these two orthogonal polarizations. By extending the calculations of equations (5) to include the effects of the quarter wave plate and the rotating half waveplate of Figure 3 the responses to vertically and horizontally polarized light may be determined. For a vertically polarized input, the vertical and horizontal components of the output are, respectively:

$$A_V = \cos\left(\frac{2\pi c}{\lambda} t + \frac{\Delta\phi}{2}\right)$$
$$A_H = 0 \ . \tag{15}$$

Similarly, for a horizontally polarized input, the output components are

$$A_V = 0$$
$$A_H = \cos\left(\frac{2\pi c}{\lambda} t + \frac{\Delta\phi}{2} - \frac{3\pi}{2}\right) . \tag{16}$$

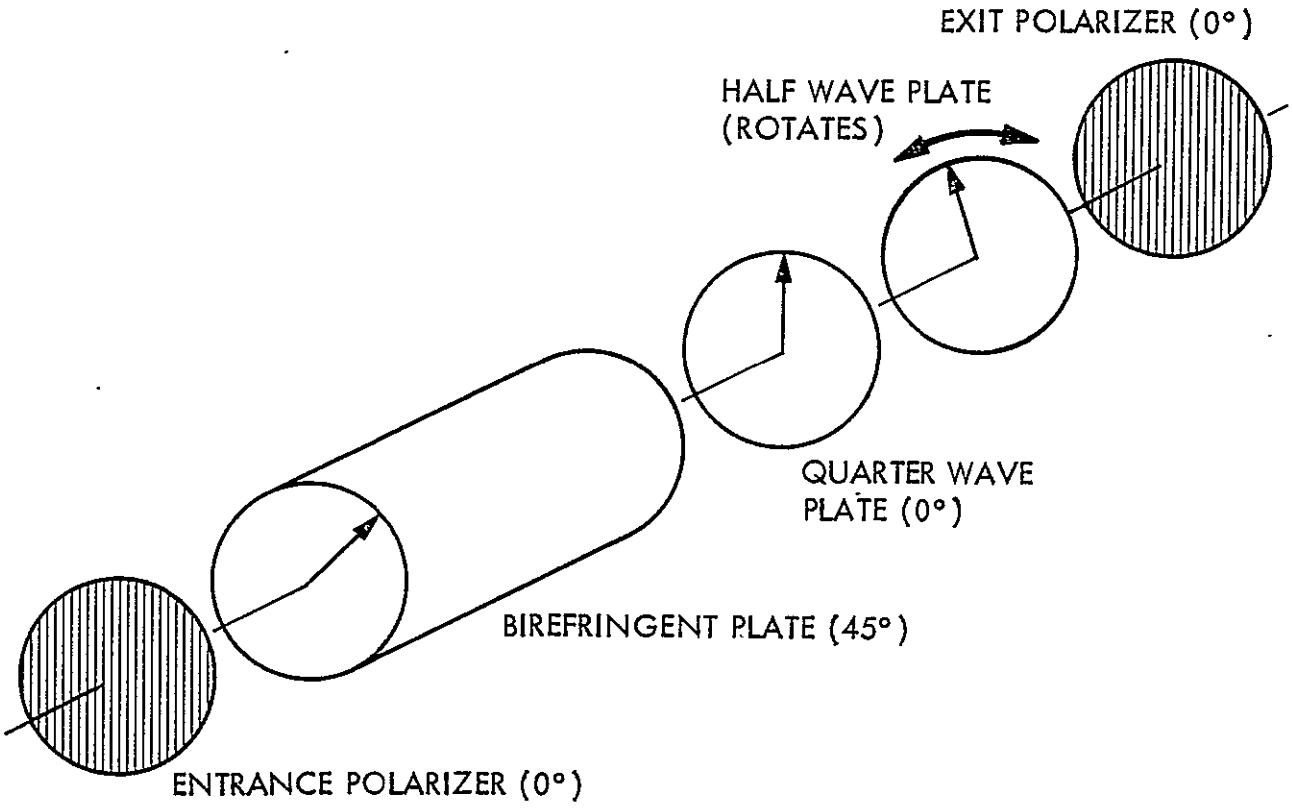


Figure 3. Diagram of a tuning element with a rotating half wave plate tuner.

Equations (15) and (16) assume that the system is tuned to the wavelength, λ, by rotating the half waveplate by an angle

$$\Theta = \frac{\Delta\phi}{4} = \frac{\pi c \Delta t}{2\lambda} \quad . \tag{17}$$

An examination of (15) and (16) indicates that vertically or horizontally polarized input becomes vertically or horizontally polarized output respectively. This is what is required, except for the $3\pi/2$ phase lag indicated by (17). This phase lag implies that arbitrary linear input polarizations will become elliptically polarized outputs. The addition of an additional quarter waveplate, oriented as shown in Figure 4, will solve this problem by introducing an extra $\pi/2$ phase lag in the horizontal component. The resulting configuration becomes a tunable element which may be used in Solc or partial polarizer filters. With the addition of entrance and exit polarizers it will still work in Lyot filters, although the extra quarter waveplate is then unnecessary.

There are filter configurations for which modifications must be made to the above tuning system because of the introduction of waveplates in the filter for other purposes. These usually only involve a fixed rotation or the addition of achromatic waveplates. The principles described above, however, remain the same.

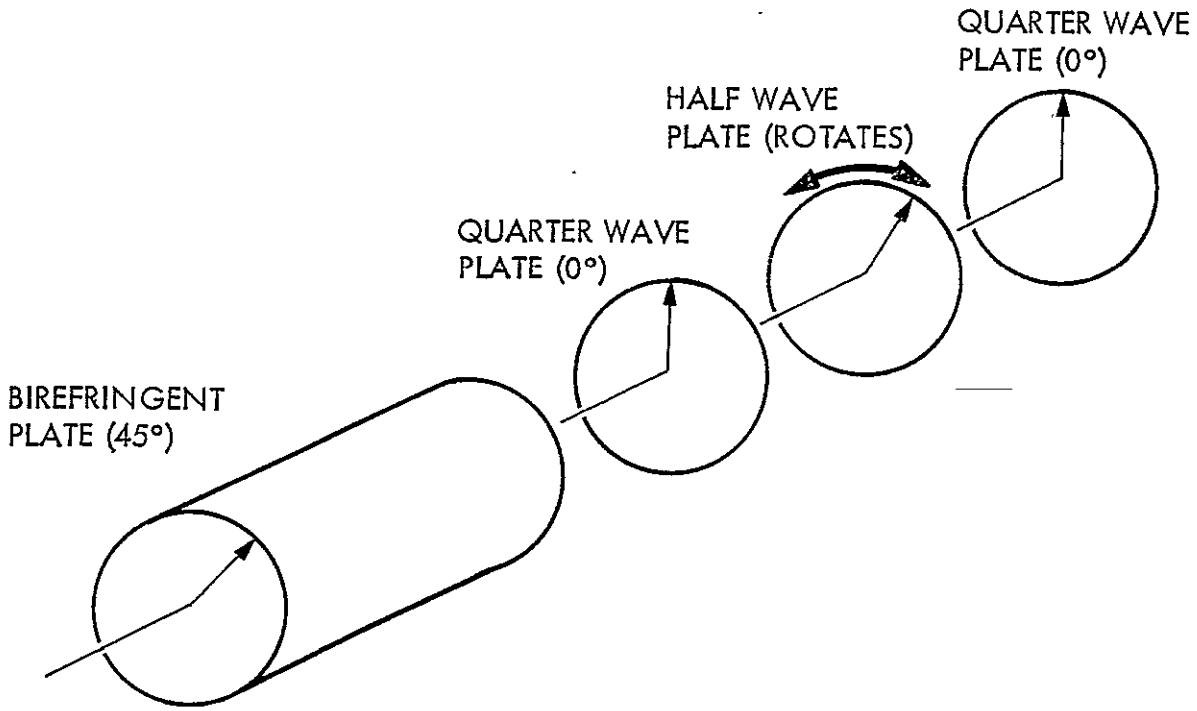


Figure 4. Diagram of wave plate arrangement for tuning a Solc or partial polarizer element.

(5) Field of View

One of the most significant advantages of the birefringent filter is its large field of view compared to Michelson and Fabry-Perot interferometric filters. This advantage, which is between 50 and 300 means that a birefringent filter can gather 50 to 300 times more light than a comparable diameter interference filter. In order to achieve this field of view advantage, the elements of the birefringent filter must be constructed as widefield elements. The widefield construction was discovered by Lyot and was subsequently described by Evans. Recently, it was discovered that special materials and alignment perturbations can increase the field by another factor of 3 to 10. A complete analysis of these field of view effects appears in Title and Rosenberg (1979) which is included as an appendix to this report. In this section a brief analysis will be presented accompanied by a summary of some of the important results.

A widefield birefringent element is made by dividing an ordinary element in half, rotating the second half around the optical path by 90^{o} with respect to the first half, and then separating the halves by an achromatic half waveplate. The half waveplate is oriented parallel to the entrance and exit polarizers, at an angle of 45^{o} with respect to the first crystal. As discussed in the preceding section, a half waveplate rotates a plane of polarization by twice the angle between the input plane and the axis of the waveplate. An impulse exiting the first crystal along the fast axis, at 0^{o}, will be rotated by the waveplate to an orientation of 90^{o}. It will, therefore, also traverse the fast axis of the second crystal. Similarly, a pulse along the slow axis of the first crystal will also traverse the slow axis of the second crystal. For a normally incident light ray the widefield element will, therefore, have the same retardation as a simple, narrow field, element.

The reason that simple elements aren't wide field is that, for non-normally incident rays, the retardation of a plate varies with the angle of incidence, i, as strongly as in a Fabry-Perot. Unlike a Fabry-Perot, however, it also varies with the azimuthal orientation, θ, of the ray. In particular for rays at 0^{o} azimuth, with respect to the

crystal's optic axis direction, the retardation decreases with increasing incidence angle, i as:

$$R(i,0^o) = R_o\left[1 - \frac{\sin^2 i}{2n_o^2}\right] . \tag{1}$$

At an azimuth of 90^o the retardation increases with incidence angle as:

$$R(i,90^o) = R_o\left[1 + \frac{\sin^2 i}{2n_o n_e}\right] \tag{2}$$

where n_o and n_e are the ordinary and extraordinary indices of refraction, respectively, and R_o is the nominal retardation of the plate. For arbitrary azimuths the retardation is given, as a second order approximation, as:

$$R = R_o\left[1 - \frac{\sin^2 i}{2n_o^2}\left(\cos^2\Theta - p \sin^2\Theta\right)\right] \tag{3}$$

$$\text{with } p = \left(\frac{n_o}{n_e}\right) .$$

The azimuthal behavior is the key to making a widefield element. In the quadrant containing the optic axis, $\Theta = 0^o$, the retardation variation is negative with incidence angle, as in (1), while in the next quadrant it is positive, as in (2). Thus, in the widefield element with the crystals orthogonal, these positive and negative quadrants overlap each other. An off-axis ray at 0^o azimuth, as it goes through the first crystal will have less retardation than a normally incident ray. As it continues through the second crystal it will have a relative azimuth of 90^o, since the crystal is rotated and therefore have more retardation than a normally incident ray. Overall, therefore, the retardation variation will be reduced as these positive and negative quadrants tend to cancel. To find out what the residual variation is equation (3) can be used twice, once for each crystal. The result is

$$R_{WF}(i,\Theta) = \frac{R_o}{2}\left[1 - \frac{\sin^2 i}{2n_o^2}\left(\cos^2\Theta - p\sin^2\Theta\right)\right]$$

$$+ \frac{R_o}{2}\left[1 - \frac{\sin^2 i}{2n_o^2}\left(\cos^2(\Theta+90) - p\sin^2(\Theta+90)\right)\right] \qquad (4)$$

which simplifies to

$$R_{WF}(i,\Theta) = R_o\left[1 - \frac{\sin^2 i}{4n_o^2}\,\frac{n_e - n_o}{n_e}\right]$$

$$= R_o\left[1 - \frac{\sin^2 i}{4n_o^2}(1-p)\right]. \qquad (5)$$

There is still some retardation variation with, i, but it is considerably reduced from (3). The factor (1-p) is very small for all known birefringent materials, always less than 0.2. For example, for calcite it is .11 and for quartz it is .006. Thus, for calcite there is about a factor of 20 difference between the coefficients of $\sin^2 i$ in (3) and (5). For even smaller index differences, the angular retardation variation decreases even further, but for very small differences the second order approximations (3) and (5), are no longer adequate.

The actual expression for the single crystal retardation is:

$$R(i,\Theta) = \frac{d}{c}\left\{n_e\left[1 - \left(\frac{a^2}{n_o^2} + \frac{b^2}{n_e^2}\right)\right]^{1/2} - n_o\left[1 - \left(\frac{a^2+b^2}{n_o^2}\right)\right]^{1/2}\right\} \qquad (6)$$

in which

$a = \sin i \sin \Theta$

$b = \sin i \cos \Theta$

d = crystal length

and the nominal retardation is

$$R_o = \frac{d}{c} (n_e - n_o) .$$

Though this is difficult to work with, increasingly higher order approximations have been developed. These become important for small index differences or relatively large bandwidth elements.

The relative variation in transmitted wavelength is the same as the relative variation in retardation,

$$\frac{\delta\lambda}{\lambda_o} = \frac{\delta R}{R_o} . \tag{7}$$

When the filter is used in an optimal optical system the field of view is set such that the wavelength shift of the extreme rays is equal to the FWHM at normal incidence. Thus, the field of view is limited by

$$\left.\frac{\delta\lambda}{\lambda_o}\right|_{max} = \frac{FWHM}{\lambda_o} . \tag{8}$$

Equation (5) is applicable whenever

$$\frac{FWHM}{\lambda_o} < \frac{(n_e - n_o)^2}{2n_e^2} . \tag{9}$$

For smaller index differences or longer bandwidths the higher order terms must be used. This is illustrated in Figure 1 in which the incident angle limits are shown for quartz and calcite, as a function of the bandwidth or relative wavelength tolerance. In Figure 1, the bandwidth shown is the equivalent bandwidth when the element is used in a telecentric optical system since this is the usual application. A telecentric system insures that the rays have the same angular distribution at all points within the

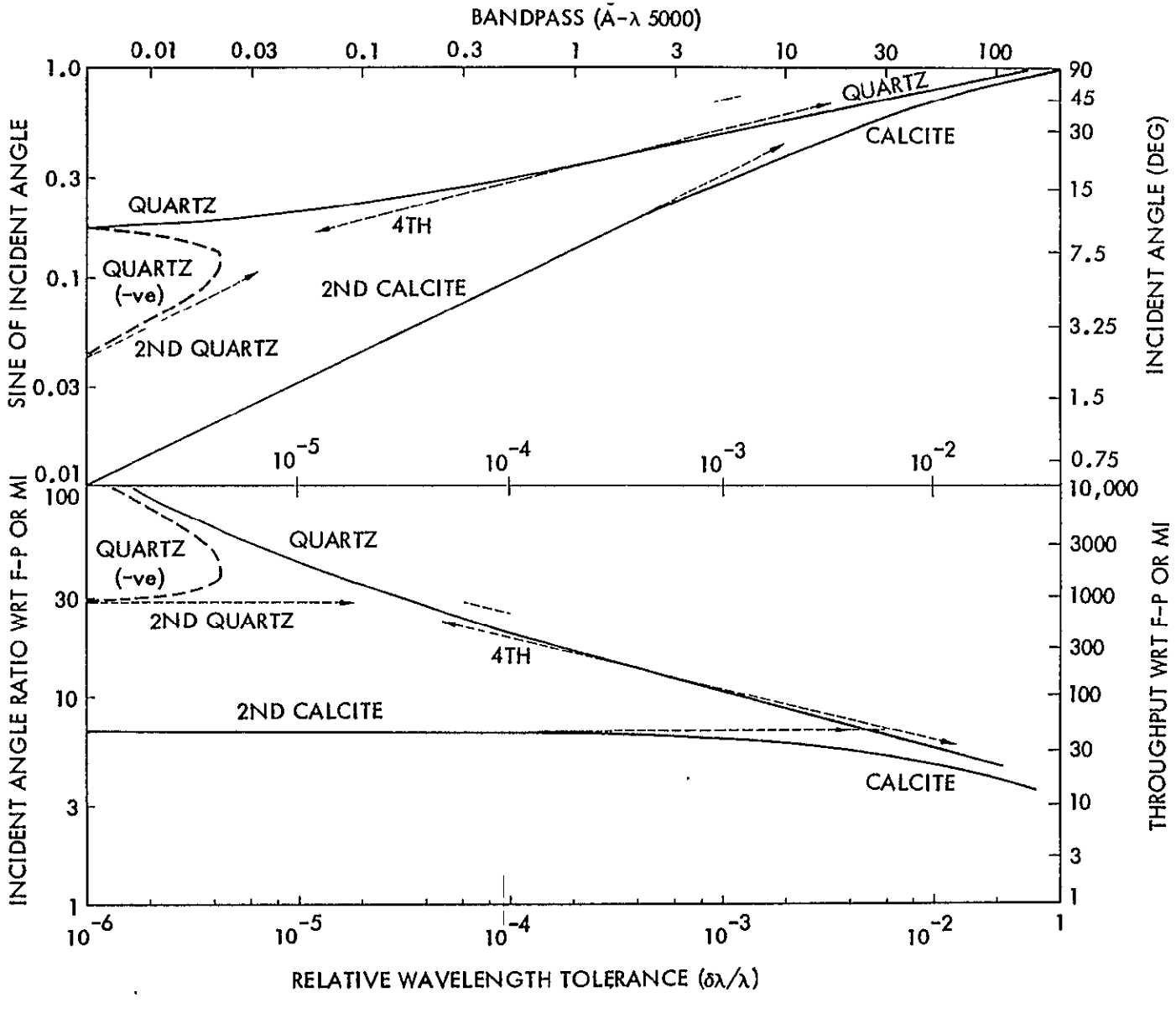


Figure 1. Incident angle and sine incident angle versus bandpass and relative wavelength shift in a cone of that incident angle (top) for quartz and calcite. Also shown is the incident angle ratio and throughput ratio with respect to F-P and MI versus bandpass and relative wavelength shift (bottom). On both halves of the figure the 2nd and 4th order approximation are shown dashed.

field of view, but at the expense of increasing the filter bandwidth by a factor of $\sqrt{2}$ over collimated, normally incident rays.

It is well known that for air spaced Michelson or Fabry-Perot filters the relative wavelength tolerance is

$$\frac{\delta\lambda}{\lambda_o} = -\frac{\sin^2 i}{2} . \tag{10}$$

If (10) is compared with (5) it is seen that, for equal wavelength tolerances, there is a solid angle, or throughput, advantage for birefringent filters of:

$$\frac{\sin^2 i_{BR}}{\sin^2 i_{FP}} = \frac{2n_o^2}{1-p} \tag{11}$$

This throughput advantage, and the corresponding incident angle ratio, is also shown in Figure 1 as a function of the telecentric bandwidth.

Besides the throughput per unit aperture advantage of birefringent filters, there is also a significant potential advantage in aperture size. Currently calcite is available in crystals of 40 mm to 50 mm diameter, but quartz can be obtained with diameters up to 300 mm. Even more exciting is the possibility of building filters using oriented plastic film as a birefringent material. These could have apertures as large as 1 meter diameter.

Table I summarizes some of the properties of birefringent elements. The index difference, throughput advantage, and FWHM-length product are shown for various materials. Two plastics are included, PVA, polyvinyl alcohol, and PET, poly (ethylene terephthalate). PET is more commonly known by the duPont tradename, Mylar. As the table indicates, the birefringence of plastic film is not unique and depends on the degree to which the material has been oriented.

Table I

Material	$(n_o - n_e)$	Throughput Advantage*	FWHM-length**
Calcite	.1720	47.5	.727
KDP	.0454	151.6	2.76
MgF	.0118	450.5	10.58
Quartz	.0091	813.	13.7
PVA	.003-.013	370.-1500.	10.-42.
PET	.05-.20	25.-98.	2.5-6.

* Throughput advantage = $\left|\frac{2n_o^2 n_e}{n_e - n_o}\right|$, and is relative to an air spaced Michelson or Fabry-Perot filter.

** FWHM-length is in units of $\mathring{A}$-cm (10^{-12} m^2), and is measured at $\lambda_o = 5000\ \mathring{A}$.

(6) Bandpass Control

In a preceding chapter the question of profile management and bandwidth determination was discussed, but all the alternatives offered there were design decisions before the filter was constructed. None, except perhaps rotating all the crystal elements of a Solc type filter, had any capability to adjust the bandwidth after the filter was constructed. There are, however, applications for which the ability to control the bandwidth of a filter is useful. As the bandwidth is increased the transmission increases, so if there isn't a need for particularly fine spectral resolution a wider bandwidth could be desirable. A design for which the bandwidth may be adjusted in use would, therefore, be advantageous.

For a Lyot filter, the bandwidth can be adjusted by a physical or optical removal of the longest element. That would double the bandwidth without changing the free spectral range. Physical removal, while conceptually simple, has many practical difficulties and is considered an unacceptable approach. Optical removal is possible and the rest of this section will discuss several, both simple and sophisticated, methods which are feasible.

If the filter is a narrow field filter, then each element is made from a single crystal (with, perhaps, a tuning section on its end). Optical removal of such an element is enabled by using two crystals, each half as long as the original. Ordinarily the two halves would be aligned with their fast axes parallel. Their retardations will, of course, add and the element will be active in the filter. To remove the element one half is rotated by 90^{o}. The retardations will subtract and the element will effectively be removed. This has, however, two disadvantages. The first is that total removal is wasteful. It would be more useful if only part of the retardation is removed, thus enabling the element to be used as a contrast element for profile control in the wider bandwidth mode. This problem could be remedied by splitting the element into non-equal pieces, which would make the element effectively short but non-zero in the rotated configuration. The second problem is that, as discussed previously, rotating a large retardation crystal element is undesirable.

Fortunately, as with a tuning section, achromatic quarterwave and half waveplates may be used to accomplish the rotation. If a half waveplate, oriented at 0°, is placed between two crystals which are oriented at $+45^{\circ}$ and -45°, respectively, then the retardations will add. Now, replace the half waveplate by two quarter waveplates, each oriented at 0° and nothing changes since they add to equal a half waveplate. If, now, one of the quarter waveplates is rotated to 90°, the quarter waveplates cancel, and effectively disappear from the optical path. The two crystals, being orthogonal, also subtract retardation. By the rotation of a single quarter waveplate, therefore, the retardation of an element may be switched from long to short. Similarly, if the crystal is divided into several pieces, say of lengths 1, 1, 2, 4, 8, ... , separated by quarter waveplate pairs, then the element can be adjusted to synthesize any even length by appropriate waveplate rotations.

In practice, however, the situation is almost as simple, but the theory must be extended to the wide field elements described in the previous section. Remember that a widefield element is constructed by splitting an ordinary element and placing a half waveplate between the two orthogonally oriented halves. The obvious realization of a variable bandwidth widefield element is, therefore, just two variable bandwidth ordinary elements, orthogonally oriented, separated by a half waveplate. As long as the two halves are each adjusted to the same retardation the combination will behave as a widefield element of twice the retardation. Thus, to make a widefield element which can be switched from a retardation of 16 to a retardation of 4 the following configuration would suffice:

$$5(0^{\circ}):Q(45^{\circ}):Q(\pm 45^{\circ}):3(0^{\circ}):H(45^{\circ}):3(90^{\circ}):Q(\mp 45^{\circ}):Q(-45^{\circ}):5(90^{\circ}) \qquad (1)$$

where Q and H denote quarter and half waveplates, respectively, and the orientations of each plate are in parentheses. Such an element could be switched between a narrow bandwidth or a contrast element role in a Lyot filter

The configuration shown above may be simplified slightly by replacing each quarter waveplate pair with a single rotatable half waveplate. This doesn't work for a single narrow field element since the extra half wave

of retardation will shift the transmission, but in the split widefield element the two outer half waveplates will cancel. The resulting configuration will then be:

$$5(0^{\circ}):H(\begin{smallmatrix}0^{\circ}\\45^{\circ}\end{smallmatrix}):3(0^{\circ}):H(45^{\circ}):3(90^{\circ}):H(\begin{smallmatrix}0^{\circ}\\45^{\circ}\end{smallmatrix}):5(90^{\circ}) . \qquad (2)$$

An important feature of this configuration is that its field of view properties correspond to its effective retardation, not its total crystal length. When it is adjusted to have retardation 16, it has the field of view of a widefield element of retardation 16. After rotation of the two outer half waveplates, however, it has both the retardation and the wider field of view of a widefield element of retardation 4. Though this property is not obvious it will become more apparent in the subsequent discussion.

Besides the variable bandwidth feature there is another reason to consider configurations such as the example above. In order to make the shortest elements of a Lyot filter, very thin crystal plates must be fabricated. Especially in widefield elements, where each plate is only half its normal thickness, this becomes a very difficult procedure. Additionally, the resulting plates are very fragile and impractical to use in any system which is subject to mechanical shock. For this reason it is desirable to use thicker crystals in the subtractive mode described above. For example, a widefield element of retardation 2 could be constructed from longer crystals as:

$$5(0^{\circ}):4(90^{\circ}):H(45^{\circ}):4(0^{\circ}):5(90^{\circ}) . \qquad (3)$$

The remarkable property of configuration (3) is that it has the full field of view of a widefield element of retardation 2. It is, therefore, a true equivalent of such an element.

There is a moderately simple extension of the mathematical analysis of widefield elements which allows for the analysis of more complicated crystal/waveplate configurations. In particular it specifies easily verifiable conditions for which a configuration is widefield. From the discussion in the Field of View section, the off-axis behavior of a single crystal is:

$$R(i,\Theta) = R_o \{1 - \frac{\sin^2 i}{2n_o^2} [1 - (1 + p)\sin^2\Theta]\} \quad (4)$$

where $R(i,\Theta)$ is the effective retardation at the angle of incidence, i, and azimuth, Θ, and p is the ratio of the ordinary to the extraordinary indices of refraction:

$$p = n_o/n_e \, . \quad (5)$$

For the widefield element:

$$R_o(0^o):H(45^o):R_o(90^o) \quad (6)$$

the retardation of the two crystals add and the field behavior becomes:

$$R_{WF} = R(i,\Theta) + R(i,\Theta + 90^o) \quad (7)$$

$$= R_o \{2 - \frac{\sin^2 i}{2n_o^2} [2 - (1+p)\sin^2\Theta - (1+p)\cos^2\Theta]\}$$

$$= 2R_o - \frac{\sin^2 i}{2n_o^2} [2R_o - R_o(1+p)\sin^2\Theta - R_o(1+p)\cos^2\Theta] \, . \quad (8)$$

Since the coefficients of the $\sin^2\Theta$ and $\cos^2\Theta$ terms are equal, the azimuthal dependence vanishes and (8) becomes:

$$R_{WF} = 2R_o - R_o \frac{\sin^2 i}{2n_o^2} [1-p] \, . \quad (9)$$

The vanishing of the azimuthal dependence is equivalent to the widening of the field since p is nearly unity. In the following, the coefficients of $(1+p)\sin^2\Theta$ and $(1+p)\cos^2\Theta$ will be denoted by S and C respectively. When S equals C the element will be widefield.

Suppose that the first half of the widefield element is split into two parts of lengths aR_o and $(1-a)R_o$, and that the second part is oriented

at 90^{o}. The retardations of these two pieces subtract and the field behavior of the whole element becomes:

$$R = aR(i,\theta) - (1-a)R(i,\theta+90) + R(i,\theta+90) \tag{10}$$

$$= a[R(i,\theta) + R(i,\theta+90)]$$

$$= 2aR_o - aR_o \frac{\sin^2 i}{2n_o^2}(1-p) \; . \tag{11}$$

This is still a widefield element, but with retardation, $2aR_o$. The field behavior is determined by the coefficient of

$$\frac{\sin^2 i}{2n_o^2}(1-p)$$

and is easily seen to be that of an ordinary widefield element of retardation $2aR_o$.

It may be easily verified that for any division of the two crystals the field behavior is given by an expression similar to (11). All that is required is that the sum of the crystal lengths on one side of the half waveplate be equal to the sum on the other side, assuming that all the crystals are oriented at $\pm 45^{o}$ with respect to the waveplate. If this is satisfied, then the element will be widefield.

It is relatively easy to generalize this result to configurations with more than one half waveplate. Assuming that all the waveplates are at 45^{o}, and that the crystals are at 0^{o} or 90^{o} the following rules characterize the behavior of a configuration:

(1) Crystal at 0^{o}

If the crystal follows an even number of half waveplates it contributes positively to the retardation.

If the crystal follows an odd number of waveplates it contributes negatively to the retardation.

The azimuthal contribution is to the coefficient, S, and is equal in magnitude and sign to its retardation contribution.

(2) Crystal at 90°

If the crystal follows an even number of waveplates its retardation contribution is negative.

If the crystal follows an odd number of waveplates its retardation contribution is positive.

The azimuthal contribution is to the coefficient, C, and is equal in magnitude and sign to its retardation contribution.

If, after applying these rules to all the crystals in the element, the coefficients, S and C, are equal the element is widefield. Moreover, the field characteristics will be that of an equivalent widefield element of the same effective retardation.

As an example of the application of these rules, consider the configuration:

$$L_1(0^{\circ}):H(45^{\circ}):L_2(0^{\circ}):H(45^{\circ}):L_3(90^{\circ}):H(45^{\circ}):L_4(90^{\circ}) . \tag{12}$$

The contributions to the retardation and the azimuthal coefficients are, respectively:

$$\begin{aligned} R &= L_1 - L_2 - L_3 + L_4 \\ S &= L_1 - L_2 \\ C &= -L_3 + L_4 . \end{aligned} \tag{13}$$

The widefield condition is, therefore,

$$L_1 - L_2 = L_4 - L_3 \tag{14}$$

and the retardation is, in the widefield mode,

$$2(L_1 - L_2). \tag{15}$$

The two outermost waveplates in (12) may be optically removed by rotating them to 0°, as in (3), which results in:

$$L_1(0^{\circ}):L_2(0^{\circ}):H(45^{\circ}):L_3(90^{\circ}):L_4(90^{\circ}) . \tag{16}$$

The respective contributions are:

$$\begin{aligned} R &= L_1 + L_2 + L_3 + L_4 \\ S &= L_1 + L_2 \\ C &= L_3 + L_4 . \end{aligned} \tag{17}$$

For this configuration, the widefield condition is:

$$L_1 + L_2 = L_3 + L_4 . \tag{18}$$

In order to use this element as an adjustable element, both (14) and (18) must be satisfied simultaneously. These imply

$$\begin{aligned} L_1 &= L_4 \\ L_2 &= L_3 \end{aligned} \tag{19}$$

with the retardation switchable between

$$\begin{aligned} R &= 2(L_1 - L_2) \\ \text{and} \quad R &= 2(L_1 + L_2) . \end{aligned} \tag{20}$$

This is the same result as in (3), but now it should be more apparent why the field behavior is the same as the equivalent widefield elements.

(7) Engineering Considerations

The previous chapters have been largely theoretical and have examined the optical principles of birefringent filter systems. This theoretical development, however, has largely been guided by a desire to design operational filter systems. It has been the goal of the entire filter study program to construct filters which are suitable for operational ground or space telescopes. Consequently, the concepts of the previous chapters are the theoretical foundations for systems which actually work.

There are some engineering problems which are encountered in the reduction to practice of birefringent filters. Some of these are typical of optical systems, and others are peculiar to birefringent filters. Generally the engineering difficulties may be partitioned into three categories: mechanical, thermal, and electronic. While none of these present serious engineering obstacles to filter construction and operation, they represent legitimate considerations.

The primary mechanical problem is mounting the crystals sufficiently firmly for the optical requirements subject to the constraint that they are not fractured by shock or acoustic loading. A method of solving the mechanical mounting problem is described in detail in Appendix 2, Development of Filters for Spaceflight. The basic concept is the mounting of the crystals in a primary holder with a thin non-hardening adhesive such as silicon rubber. These primary holders are then mounted in a fixed (non-rotating) structure which serves to provide both mechanical alignment and good thermal contact to the crystals. As stated previously, both spectral and image motion problems occur when crystals are rotated, so for high quality imaging systems the crystals are fixed, and only thin, quarter and half wave plates are rotated.

Before the development of inexpensive electronic computers, tuning narrow band filters presented a severe mechanical problem. All tuning elements of the filter were geared together and rotated by a single drive. Since these were Lyot filters with 2:1 crystal length ratios, each element had to rotate twice as far as its predecessor. Additionally, the optical lengths of the crystals had to be very precisely 2:1 or the system would quickly lose spectral alignment during tuning. These restrictions are

so severe that mechanically geared, narrow bandwidth filters never operated over more than ±20 Å without some differential corrections to the element cascade.

With presently available minicomputers, however, these problems are easily solved. It is straightforward to tabulate, or have an accurate analytic expression for, all the transmission peaks in the channel spectrum of each filter element. Each element may then be tuned individually by rotation of its tuning waveplate, under computer control. To tune an element through its free spectral range requires a 180° rotation, implying a wavelength accuracy of FSR/180° per degree of rotational accuracy. Since the free spectral range of each element is just twice its full width at half maximum, a wavelength accuracy of FWHM/100 implies a rotational accuracy requirement of only 0.9° for even the thickest element. This is well within the capabilities of even modest shaft encoders.

The rotational accuracy requirements for birefringent filters should be compared with the equivalent requirements for Fabry-Perot or Michelson interference filters. There, the path lengths must be adjusted to accuracies of small fractions of wavelengths. The rotational requirements of tunable birefringent filters represent very simple mechanical problems by comparison.

Thermal control of birefringent filters has been a serious problem in the past, but, like tuning, it is solvable with minicomputer control. The temperature sensitivity of a 0.1 Å FWHM calcite element is about 0.3 Å/°C which means that holding the element to an accuracy of FWHM/10 implies a temperature control to .03°C. While this has been an extremely difficult requirement previously, it may be solved by using the filter's own tuning capability. Instead of trying to maintain a constant thermal environment to .03°C, it is easier to monitor the temperature and compensate for variations by retuning each element. This requires precise knowledge of the temperature sensitivity of the filter materials, good temperature sensors, good thermal design of the filter housing and a small minicomputer to do the necessary calculations. Fortunately, none of these represent severe restrictions. The filter itself can be used to determine the temperature sensitivity of the materials by placing the

filter assembly in an oven and monitoring the spectral profile at various temperatures. The required sensors and minicomputer are now readily available. Two 0.1 Å FWHM filters have been built using the temperature compensation by tuning concept and they have proven to be accurate to .01 Å over a temperature range of 40°C.

The filter electronics package serves as an interface between the filter and the minicomputer. It must be able to operate the drive motors for the tuning sections and also monitor the temperature sensors. The electronics package is well within the state of the art and represents no serious problems.

There are, of course, problems associated with the preparation and assembly of the filter components. These filters are sophisticated optical assemblies and must be given the same care as any precise optical device. Fortunately, as in other optical construction, there are highly sensitive interferometric testing techniques which can insure proper preparation and alignment. Moreover, each component or subassembly of a birefringent filter is, itself, an interferometer. Thus each component need only be viewed in monochromatic light to produce the diagnostic fringe patterns.

The same analysis which was used to determine the field characteristics of a well-aligned wide-field element can also be used to determine the behavior of misaligned or mismatched elements. This has been done, and a catalog of fringe patterns for various states of adjustment has been prepared, and is included in Appendix 1. These representative patterns make it easy to identify and correct problems during the optical polishing and assembly stages. Each element must be prepared so that the retardation is uniform to within a fraction of a wavelength. This is not an especially difficult optical task since it is the retardation which must be uniform to $\lambda/10$, not necessarily the crystal surface. Because the birefringence is much less than unity this makes the tolerance requirement weaker than the corresponding $\lambda/10$ requirement for good optical surfaces. In particular, it is weaker by a factor of about 10 for calcite or 200 for quartz.

The distinction between optical surface quality and retardation accuracy is very important in birefringent filter construction. For a quartz element, an optical surface which is uniform to $\lambda/10$ implies a

retardation which is uniform to $\lambda/2000$. In Fabry-Perot filters, for comparison, there is no distinction. In order to achieve the required $\lambda/200$ retardation accuracy it is necessary to actually polish the surfaces to that figure. Thus, a birefringent element can achieve more precise retardation uniformity with less stringent optical polishing requirements.

Since it was recognized that there was no need to rotate the crystal elements themselves it has been possible to completely separate the preparation and assembly difficulties from the operational considerations. It has been found that a properly aligned filter, once assembled, remains stable, to the extent of being space qualifiable. This is in contrast with many optical systems in which the critical operational parameters are the same as the critical alignment adjustments. For example, in a Fabry-Perot filter a critical factor is the plate spacing, but this is just the dimension which must be changed to tune the filter. In the birefringent filters the critical alignment parameters are the crystal element orientations, but these are never changed during operation. This insures the stability of these systems.

Examples of narrow band tunable birefringent filters are shown in Figures 1 through 5. In Figure 1 all the optical components of the Lockheed Alternate Partial Polarizing Filter (LAPPU) are shown prior to assembly. The widefield elements, clearly showing the length differences, are in the background. The rotatable half wave tuning plates, the perfect internal and external polarizers and the internal partial polarizers are in the foreground. An exploded view, indicating the assembly of these components is shown in Figure 2, and the final filter with tuning motors is pictured in Figure 3. The result of evolution of the design of these filters is shown in Figure 4. This is a Spacelab 2 flight prototype filter with more compact motors and a somewhat more sophisticated mechanical drive system. Fundamentally, however, it is very similar to the LAPPU filter. Each filter has a 0.1 Å bandwidth and an aperture of 35 mm.

Figure 5 shows a view of a tunable 10 Å Lyot filter with rotating crystal elements. The aperture is 75 mm. Elements for this filter were

C-2

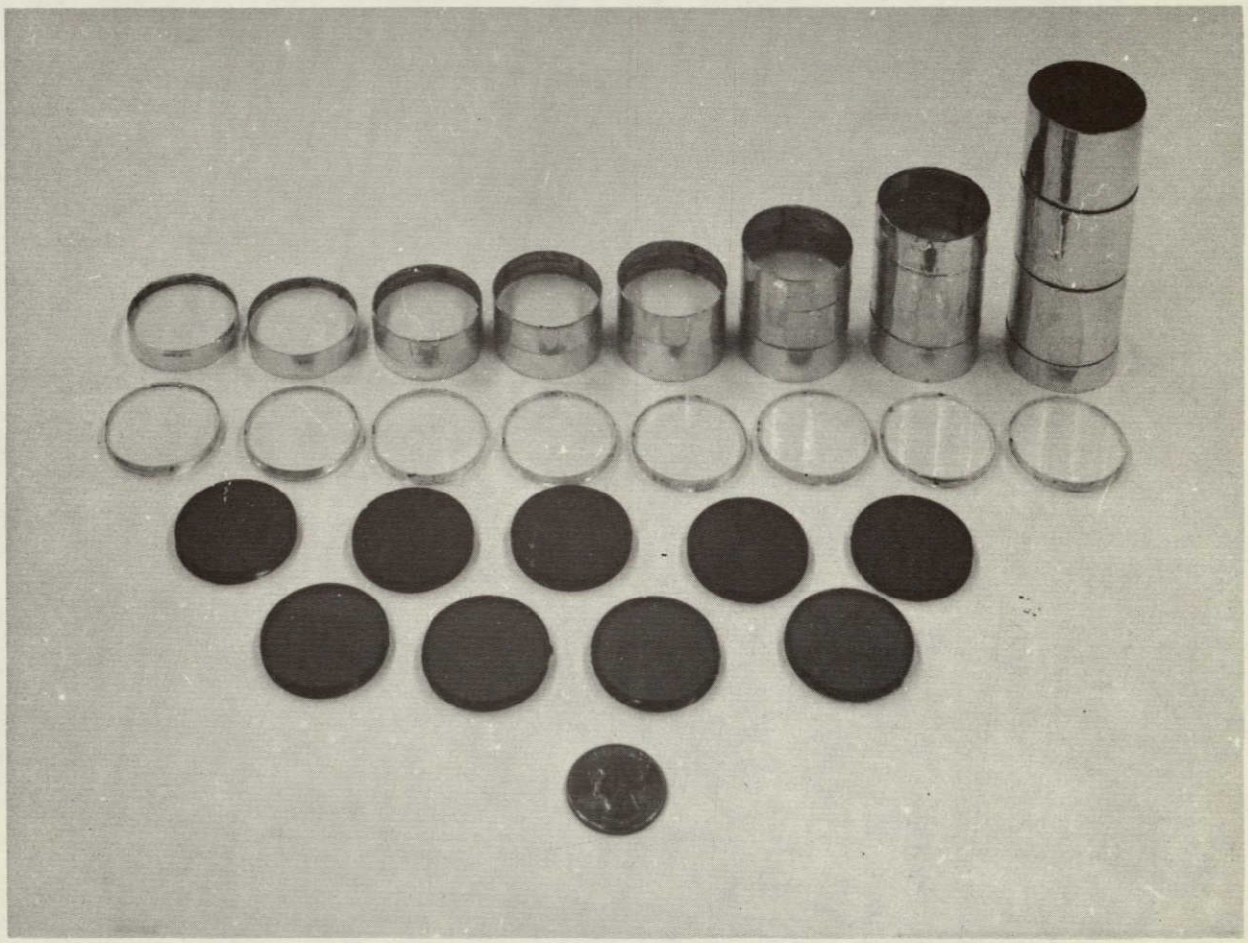


Figure 1. Optical components of the 0.1 Å LAPPU showing wide field elements, half wave plates, polarizers and partial polarizers.

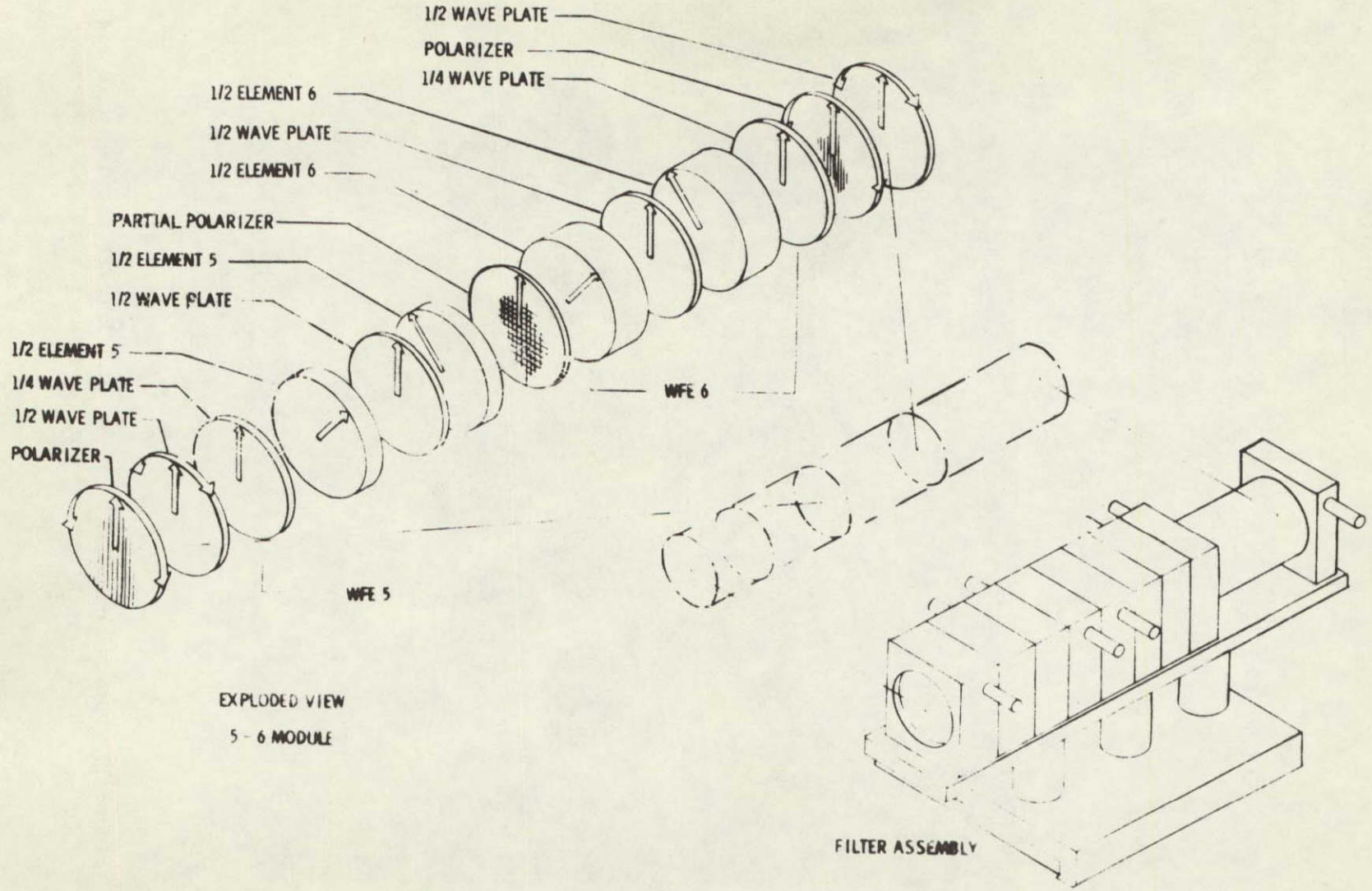


Figure 2. Exploded view of LAPPU.

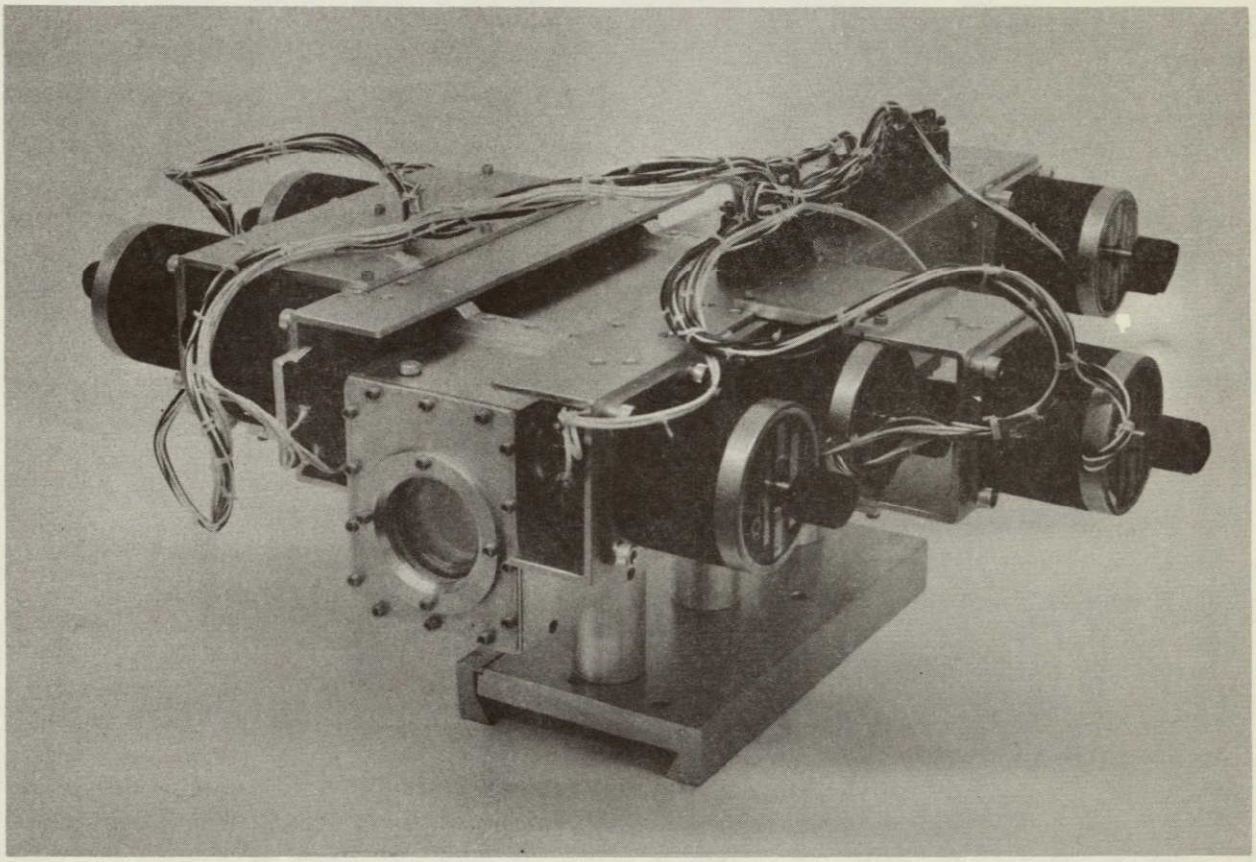


Figure 3. Complete LAPPU filter.

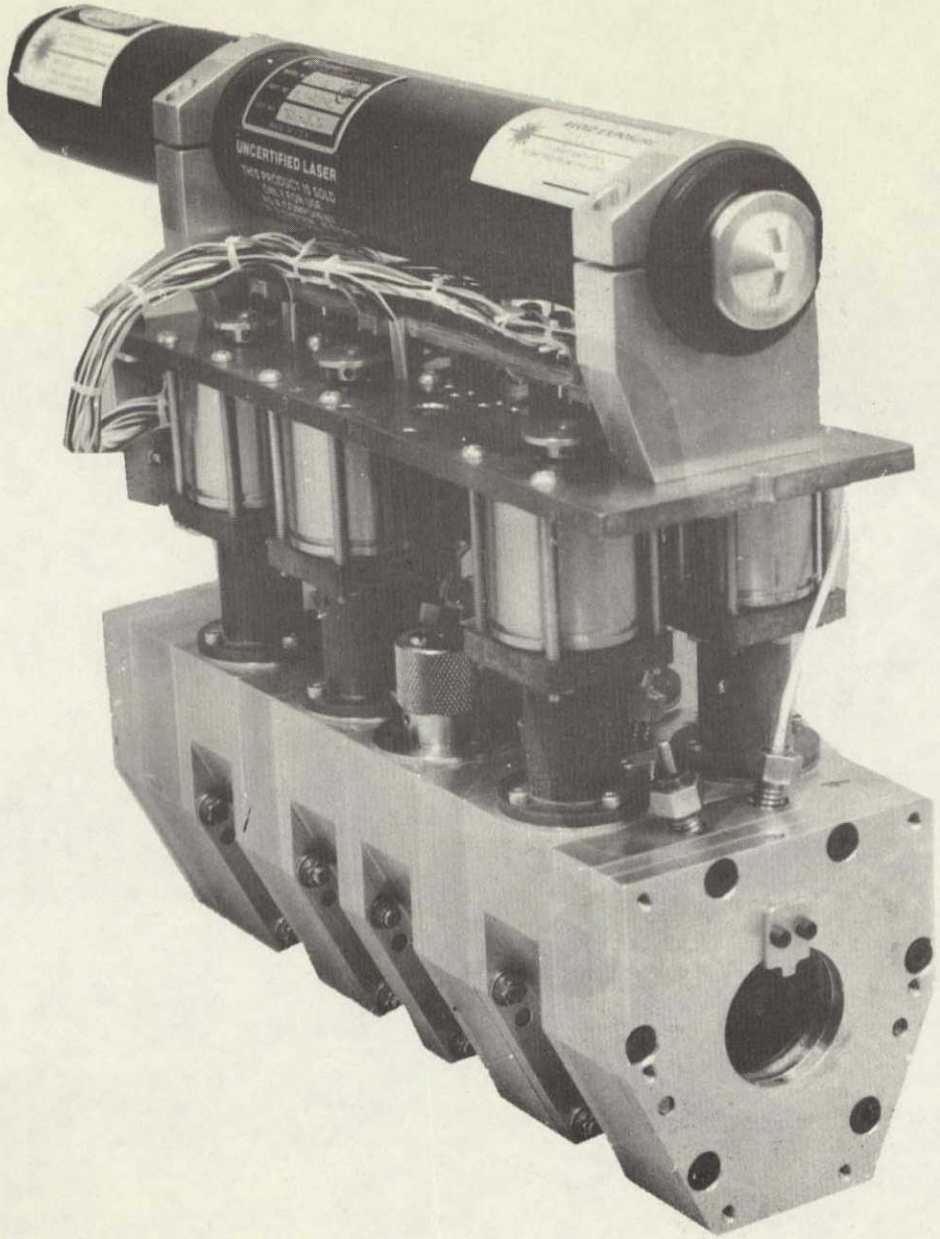


Figure 4. Flight prototype 0.1 Å filter for Spacelab 2.

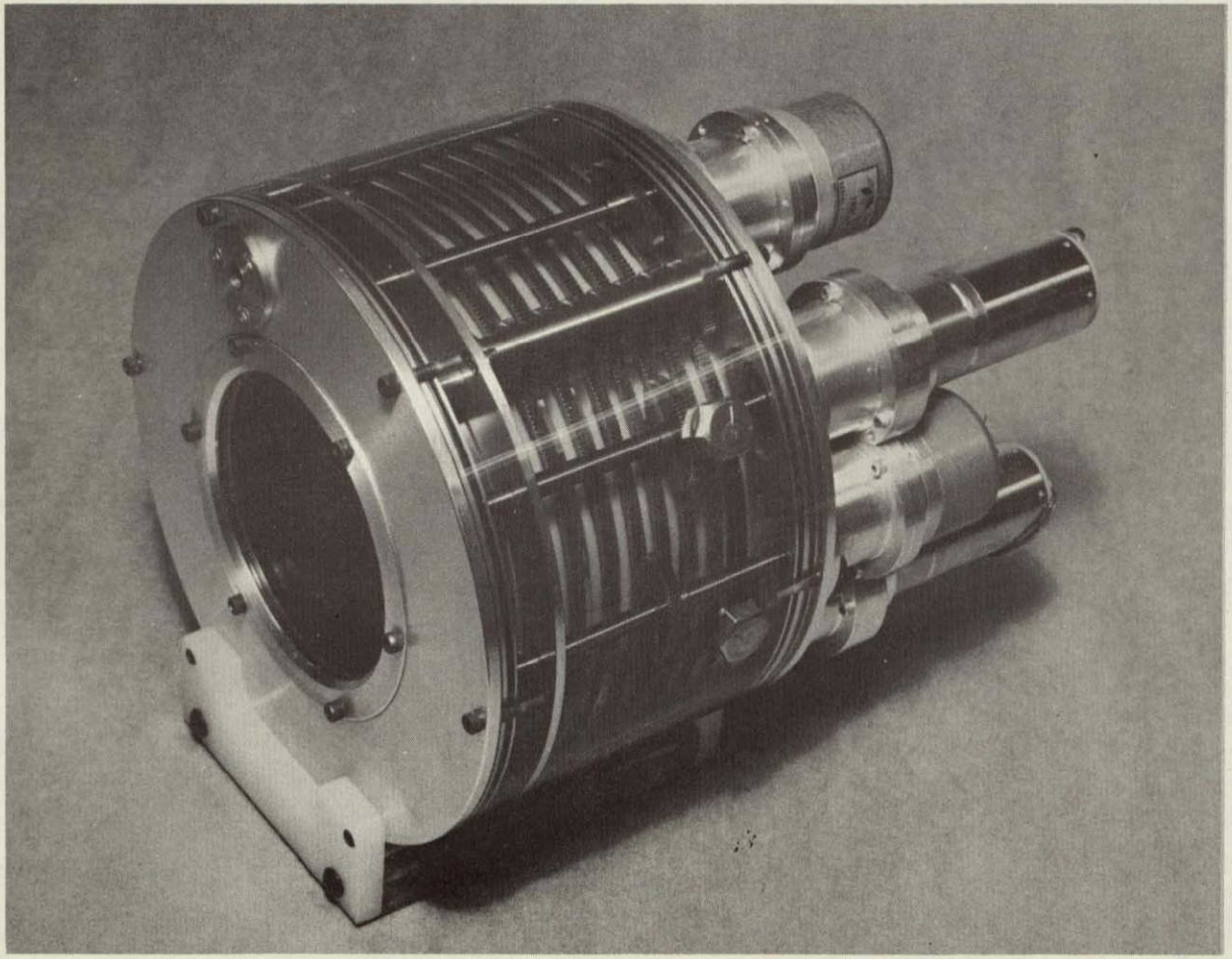


Figure 5. A 10 Å tunable filter with 75 mm clear aperture.

manufactured using techniques normally used to produce silicon wafers for integrated circuits. Although the filter operates spectrally, there are some undesirable variations introduced by the rotating elements. The filter will be rebuilt to tune by rotation of wave-plates instead. It does demonstrate, however, that direct 2:1 gearing can be used for a filter of modest bandwidth. This filter is tuned by a single motor, although two motors can be used to increase the tuning speed between widely separated wavelengths.

(8) Conclusions, Projections and Recommendations

At the start of this program it seemed clear that it would be possible to build filters, tunable through the visible spectrum, and with significant throughput gains over conventional interference filter techniques. That such filters could have bandwidths as small as .05 Å or as large as desired also seemed evident. The 0.1 Å and 10. Å tunable space flight filters which were built during this program have verified these beliefs. The theoretical and experimental results of the program, moreover, have indicated that even more may be expected.

The detailed analysis of the field of view characteristics reported in Appendix 1 has shown that an additional factor of from 3 to 10 in throughput is even possible. The field behavior predictions of this analysis have also been shown to be an effective diagnostic tool for the proper alignment of the widefield elements of the filters. Knowing the relationship between the field pattern and the alignment status makes it possible to determine the exact state of an element. This makes the actual construction of filters much more practical.

The bandpass control techniques which are discussed in Chapter 6 indicate that a single filter can be an extremely versatile instrument. Its operating characteristics may be selected, in use, to match the immediate needs of an observing program. Because an element can be configured so that it always has the field characteristics of its bandwidth equivalent, angular throughput may be traded off against spectral resolution while still exercising control over the out of band spectral profile. Thus, instead of a filter having a single operating characteristic which is matched to one scientific task it may have a selectable set of characteristics making it suitable for many tasks with no performance degradations for any one. Such versatility can be especially valuable for space platform applications.

Outside of the visible band, above 7500 Å or below 4500 Å the potential of birefringent filters wasn't as clear at the beginning of this program. The most immediate problem was the availability of suitable polarizers. In the infrared, up to about 4 μm, sheet polarizer is

available. Though similar in operation to visible wavelength polarizers its efficiency is too low to make Lyot filter designs practical. Another choice, however, now seems promising. Modern holographic and shadow deposition techniques make wire grid polarizers a real possibility down to below 1 μm and perhaps as low as 7500 Å. Several wire grid polarizers for the 1 μm region have been built and the process seems promising. An advantage to wire grid polarizers is that they have the possibility of 100% polarization efficiency.

For the shorter wavelengths, below 4500 Å, the prospect for a birefringent filter was not promising. Only crystal polarizers appeared practical because the transmission of sheet polarizers drops to below 50% very quickly as the wavelength decreases. The roughly cubical shape of the crystal polarizers makes them very unattractive for a Lyot cascade. The added length to the optical path presents serious engineering problems if many of the polarizers are used. For this reason partial polarizer and lossless, Solc, configurations were considered.

Although a 2800 Å Solc filter had been built by Fredga and Hogbom and operated successfully, there seemed to be many objections to that design. For example, Solc filters seemed difficult to tune with crystal wedges; they were difficult to manufacture because of the large number of plates; and they were difficult to align because of the sensitive, interacting plate orientations.

It was finally realized, however, that the conventional waveplate tuning techniques, discussed in Chapter 4, would apply to Solc and partial polarizing designs. This makes the Solc design seem very attractive. Similarly the silicon wafer manufacturing techniques were shown to apply to moderate aperture quartz elements by actually constructing a 10 Å tunable filter from laminated quartz plates. Most recently, moreover, it has been realized that the synthesis algorithms developed by Harris et al provide a solution to both the multitude of plates and the alignment problems.

The surprising result of Harris, namely that there are many plate orientation combinations which generate the same filter profile, offered the possibility of multiple thickness plates in some configurations. As

was shown in Chapter 6, this actually occurs and results in nearly a factor of two reduction in the number of plates. Also, as discussed in Chapter 2, the Harris synthesis procedure provides an excellent diagnostic tool for the alignment of a Solc filter. Now it appears that Solc filters will offer a practical solution for the shorter wavelength region, and for verification, an actual construction project is just beginning.

Since Solc type filters have the capability of almost arbitrary profile synthesis it is also possible to design hybrid, Solc-Polarizer-Solc, filters to achieve moderate to large finesses. The design presented in Chapter 6, while not optimal, is certainly very promising. As these designs use only one or two internal polarizers, rather than six or seven, they are still practical and make tunable ultraviolet filters a real possibility.

The filter development program has made significant advances in both the theory and practice of tunable birefringent filter design:

(1) For the visible, 4500 Å to 7500 Å, region extremely narrow band filters with transmissions exceeding 50% in polarized light have been shown to be practical and versatile. Two such filters have been qualified for space flight.

(2) In the near infrared, wire grid polarizers have been built and appear quite practical. As they offer the possibility of 100% polarization efficiency it should be possible to build practical filters in the IR. Work is progressing slowly in this area but should be increased. The Solc-type filters are also being considered for the near infrared.

(3) For the ultraviolet region between 1500 Å and 4500 Å there are two possibilities. Partial polarizing filters should work well down to 3500 Å with virtually the same techniques as were used for the longer visible wavelengths. New designs are being investigated but the principles of construction have been established. Below 3500 Å, however, the recently discovered Solc-hybrid configurations seem most promising.

Currently a crystal set is being fabricated so that experimental filter construction can begin this year.

In all, the prospects for continued development of practical, high-performance filters for space platform observations from 1500 Å to 10 μm seem very promising. It is even likely that the infrared developments may be extended to the millimeter region for which birefringent Solc filters have already been constructed.

In order to enable the continued improvement of widefield tunable birefringent filters the following development programs are recommended:

(1) Holographic and electron beam fabrication of polarizers for the infrared and ultraviolet.

(2) Artificial growth of large diameter calcite crystals.

(3) New birefringent material investigation for the various spectral regions.

(4) Electro-optical waveplate construction.

(5) Theoretical extensions to the general synthesis procedures to include hybrid and partial polarizing configurations.

and finally,

(6) The support and encouragement of non-solar applications of birefringent filters to utilize their full scientific capabilities.

Acknowledgements

The work described in this report started as a result of many pleasant discussions with Larry Mertz who introduced one of us (Title) to the utility of pulse response and time domain analysis of filters. The theoretical development has been aided and sometimes led by the many interesting and sometimes surprising experimental results obtained by Harry Ramsey. In addition to pointing out interesting observations to explain and verify predicted phenomena, Ramsey has built virtually all of the optical components of the Lockheed filters. Over the years the mechanical design and assembly of complete filters has been done by Ralph Reeves. The recent Spacelab filters were designed by Beez Rising and assembled by Reeves. Electronics for the filters have been the responsibility of Russell Lindgren. The computer software for tuning and temperature control of the filters has been designed by Steve Schoolman. Without the aid of all the above, real operational filters would not have been possible.

Over the years we have been supported emotionally and scientifically by John Evans, Dick Dunn, Keith Pierce, and Jacques Beckers. Financial support has been received from NASA, NOAA, NSF, AFGL, and Lockheed Independent Research funds. We are especially grateful to Dr. Gertz Oertel, formerly of NASA Headquarters, for encouraging and supporting the development of the LAPPU filter.

References

1. Ammann, E. O., 1965, J. Opt. Soc. Am. 55, 835.

2. Ammann, E. O., 1966, J. Opt. Soc. Am. 56, 1746.

3. Beckers, J. M. and R. B. Dunn, 1968, Report AFCRL-65-605.

4. Evans, J. W., 1949, J. Opt. Soc. Am. 39, 229.

5. Evans, J. W., 1958, J. Opt. Soc. Am. 48, 142

6. Fredga, K. and J. A. Hogbom, 1967, Report, Sonnenborgh Observatory Utrecht, Netherlands.

7. Giovanelli, R. G. and J. T. Jeffries, 1954, Austr. J. Phy. 7, 254.

8. Harris, F. J., 1978, Proc. IEEE 66, 51.

9. Harris, S. E., E. O. Ammann, and I. O. Chang, 1964, J. Opt. Soc. Am. 54, 1267.

10. Jones, R. C., 1941, J. Opt. Soc. Am. 31, 488.

11. Lyot, B., 1933, Compt. Rend. 197, 1593.

12. Lyot. B., 1944, Ann. Astrophys. 7, 31.

13. McIntyre, C. M., and S. E. Harris, 1968, J. Opt. Soc. Am. 58, 1575.

14. Ohman, Y., 1938, Nature 141, 157.

15. Pancharatnam, S., 1955a, Proc. Indian Acad. Sci. A41, 130.

16. Pancharatnam, S., 1955b, Proc. Indian Acad. Sci. A41, 137.

17. Schiffman, B. M. and L. Young, 1968, IEEE Trans Microwave Theory 6, 351.

18. Schurcliff, W. A., 1966, Polarized Light (Harvard Univ. Press, Cambridge, Mass.)

19. Slepian, D., 1978, B.S.T.J. 57, 1371.

20. Solc, I., 1953, Cassopis Pro Fysika 3, 1255.

21. Solc, I., 1958, Cassopis Pro Fysika 9, 237.

22. Solc, I., 1965, J. Opt. Soc. Am. 55, 621

23. Swindell, W., 1975, Polarized Light (John Wiley and Sons, New York)

24. Title, A. M., 1974, Solar Physics 38, 521.

25. Title, A. M., 1975a, Appl. Opt. 14, 229.

26. Title, A. M., 1975b, Appl. Opt. 14, 445.

27. Title, A. M., 1976, Appl. Opt. 15, 2871.

28. Title, A. M. and W. J. Rosenberg, 1979, Appl. Opt., in press.

Appendices

1. Improvement in Birefringent Filters V: Field of View Effects

2. Development of Birefringent Filters for Spaceflight

IMPROVEMENTS IN BIREFRINGENT FILTERS V: FIELD OF VIEW EFFECTS

A. M. Title
W. J. Rosenberg
Lockheed Palo Alto Research Labs
3251 Hanover Street
Palo Alto, CA 94304

ABSTRACT

The field of view characteristics of wide field birefringent (WFB) elements are compared to those of the Fabry-Perot (FP), Michelson (MI), and wide field Michelson (WFM) interferometers. Throughput gains of 50 to 300 or more with respect to the FP or MI are demonstrated. Further, it is shown that by proper choice of material, WFB elements can have angular characteristics identical to WFM interferometers. The properties of misaligned and mismatched half length WFB elements are calculated. It is shown that properly "Misadjusted" WFB elements can exhibit throughput gains with respect to properly adjusted systems. Finally, a catalog of fringe patterns for WFB elements of different materials and differing angular misadjustment is presented.

I. INTRODUCTION

Wide field birefringent elements offer significant advantages in field of view compared to Fabry-Perot and conventional Michelson interferometers. This advantage has been realized by solar astronomers since the invention of the wide field element by Lyot[1] in 1933. However, the relationship between wide field birefringent elements and Fabry-Perot and Michelson interferometers has not been generally appreciated. Further, although the basic approach of Lyot[2] and Evans[3] in describing the field characteristics of the wide field element was correct, the approximate expressions that they reported are only valid for a limited range of cases. In particular, they are not valid for crystals significantly less birefringent than calcite.

In the past, birefringent filters have seen use largely as solar instruments. These devices have typically operated in a single line such as hydrogen alpha and had bandpasses in the range of 1 to 0.125Å. For such systems the descriptions of Lyot and Evans are adequate; however, even in these cases the properties of elements with misaligned or imperfectly matched element half lengths cannot be properly handled with extensions of the second order expansions.

At present it appears that the significant gains in throughput, (fifty to several thousand) of birefringent elements, the easily implementable techniques for tuning over 1.25 decades in frequency and their superior out of band rejection make tunable birefringent elements very attractive systems in applications which require imaging of extended sources or light gathering from diffuse sources. In the more conventional spectrum scanning applications birefringent systems do not require the intensive data processing of Michelson systems nor do they require that the spectral region under investigation be sampled uniformly. In addition birefringent elements are not sensitive to their orientation

with respect to gravity nor are precision motion controls required for scanning. Temperature sensitivity, which used to be a problem with high resolution systems ($R>10^4$), can now easily be controlled by temperature sensing and thermal compensation via the tuning elements[4].

Because of the apparent advantages of birefringent systems it is worthwhile to establish their relation to the more conventional Fabry-Perot and Michelson interferometers. This is done in Section II. In Sections III, IV and V the formalism for the angular sensitivity of a simple birefringent plate, a well-aligned wide field element, and misaligned and mismatched wide field elements are developed. Using the formalism developed in the earlier sections the throughput advantages of the wide field birefringent element are calculated in Section VI.

The purpose of Section VII is to explain the fringe patterns produced by wide field elements of differing material properties and differing alignment states. Experimentally, it has been found that the principle areas of transmission loss in birefringent elements are due to wide field misadjustment problems. Section VII demonstrates that virtually all sources of misadjustment can be identified, understood, and corrected by reference to tne fringe patterns.

II. RELATIONSHIPS BETWEEN WIDE FIELD BIREFRINGENT ELEMENT, FABRY-PEROT, MICHELSON, AND WIDE FIELD MICHELSON INTERFEROMETERS

It is important to establish the connection between the wide field birefringent (WFB) element and the more traditional Fabry-Perot (F-P), Michelson (MI), and wide field Michelson (WFM) interferometers. This can be efficiently done in terms of the time delays introduced by the various systems. Shown in Figure 1 is an optical schematic of a simple birefringent element and a standard Michelson interferometer.

The simple birefringent element consists of an entrance polarizer, a birefringent plate with its optic axis in the plane of the crystal face and an exit polarizer. The purpose of the entrance polarizer is to prepare a single state of polarization, since light originating from different polarization states do not interfere. Waves propagating through the plate parallel to the optic axis propagate at velocity c/n_o and those perpendicular at velocity c/n_e, where n_o and n_e are the ordinary and extra-ordinary indices of refraction. Since the entrance polarizer is at 45^o to the optic axis, equal amounts of light are in the two components. After propagating along a length d of birefringent plate, the two polarized beams are separated in time by

$$\Delta_o = \frac{d}{c} (n_e - n_o), \qquad (1)$$

where Δ_o is the time delay for waves at normal incidence to the crystal. The exit polarizer being parallel to the entrance polarizer passes equal components of the two beams. Because the two beams arose from the same state and after passing through the exit polarizer they are indistinguishable, they will interfere. The intensity relation for two beam interference with time delay Δ_o is just

$$I(\lambda) = \cos^2\left(\frac{\pi \Delta_o c}{\lambda}\right), \qquad (2)$$

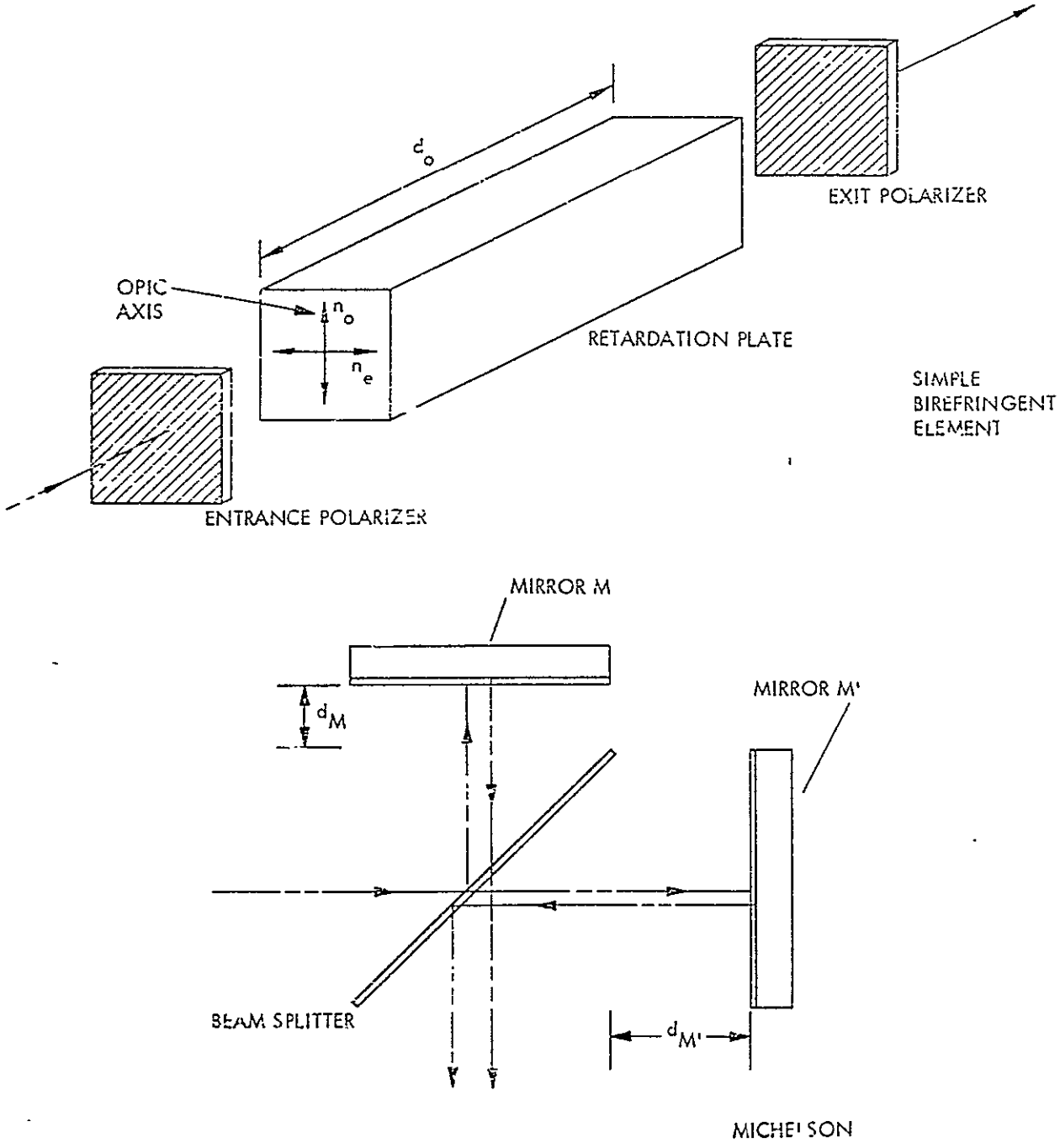


Figure 1. Optical schematic of a simple birefringent element and a normal Michelson interferometer.

where λ is the wavelength and c is the speed of light.

Now consider the Michelson; the incoming beam strikes a 50/50 beam splitter and half the beam is reflected upward while the other half is transmitted. The reflected beam strikes a mirror, M, and returns to the beam splitter. The transmitted beam strikes M' and is similarly returned to the beam splitter where half the light is reflected in the same direction as the transmitted beam from M. If M and M' are not equal distances from the beam splitter, there will be a time delay

$$\Delta_o = \frac{2}{c} (d_M - d_{M'}) \qquad (3)$$

between the two beams. Since the beams arose from the same initial states and after merging they are indistinguishable, they too will interfere. Again, the intensity relation is just that given by relation (2).

$$I(\lambda) = \cos^2 \left(\frac{\pi \Delta_o c}{\lambda}\right)$$

Note that the beam splitter functions twice in the Michelson, first to divide the incident beam into two beams, and second to recombine the two beams from the mirrors. That is, the beam splitter serves functions identical to both the entrance and exit polarizers of the birefringent element. The difference in spacing of the two Michelson mirrors has the same function as the difference in the optical path lengths along the two polarization directions in the birefringent plate.

There are, however, some angle sensitivity differences between the birefringent element and the Michelson. These arise from the off axis wave propagation characteristics of birefringent materials. For a Michelson the angular dependence of the time delay is given by[5]

$$\Delta = \Delta_o \left[1 - \left(\frac{a^2+b^2}{n^2} \right) \right]^{1/2}, \qquad (4)$$

where n is the index of refraction of the Michelson optical path, and

$$\begin{aligned} a &= \sin i \cos \theta \\ b &= \sin i \sin \theta, \end{aligned} \qquad (5)$$

where i is the angle the incident ray makes with the normal to M' and θ is the angle the projection of the ray makes with the plane of the interferometer. The theta dependence has been written implicitly to illustrate the formal similarity with the birefringent plate. From Equations (4) and (5), however, the theta dependence drops out for the MI. Most MI operate in air, but the index of the optical path has been included because the Michelson angular dependence is the same as that of a Fabry-Perot or interference filter which is often filled with a dielectric material. Equation (4) can be rewritten as

$$\Delta = \Delta_o \left[1 - \frac{\sin^2 i}{n^2} \right]^{1/2}. \qquad (6)$$

The equation is usually seen expanded to second order in sine i, or

$$\Delta = \Delta_o \left[1 - \frac{1}{2} \frac{\sin^2 i}{n^2} \right]. \qquad (7)$$

Since

$$\frac{\delta\Delta}{\Delta_o} = \frac{\delta\lambda}{\lambda_o}, \qquad (8)$$

Where $\delta\Delta$ and $\delta\lambda$ are the variations with angle of the time delay and wavelength, respectively, the relation between relative wavelength shift and incident angle

$$\frac{\delta\lambda}{\lambda_o} = -\frac{1}{2} \frac{\sin^2 i}{n^2} \qquad (9)$$

or

$$\frac{\delta\lambda}{\lambda_o} \simeq \frac{1}{2}\left(\frac{i}{n}\right)^2$$

Equation (9) is the relation normally used for evaluating F-P, MI, or interference filter systems. Neglect of the proper expression (Equation (6)) seldom causes problems.

For a birefringent element the angle dependence of the time delay[6] is

$$\Delta = \frac{d}{c}\left\{ n_e\left[1-\left(\frac{a^2}{n_o^2}+\frac{b^2}{n_e^2}\right)\right]^{1/2} - n_o\left[1-\left(\frac{a^2+b^2}{n_o^2}\right)\right]^{1/2}\right\}, \qquad (10)$$

where the theta axis coincides with the optic axis of the birefringent plate. Equation (10) represents the time difference between an ellipsoidal (extraordinary) and spherical (ordinary) wave. Expanding Equation (10) to second order, a relation similar to Equation (7) can be obtained

$$\Delta = \Delta_o\left\{1 - \frac{\sin^2 i}{2n_o^2}\left[\cos^2\theta - \left(\frac{n_o}{n_e}\right)\sin^2\theta\right]\right\}, \qquad (11)$$

where Δ_o is given by Equation (1). There are several important differences between relations (7) and (11). First, when

$$\cos^2\theta = \left(\frac{n_o}{n_e}\right)\sin^2\theta \qquad (12)$$

or

$$\tan\theta = \pm\left(\frac{n_e}{n_o}\right)^{1/2}$$

then

$$\frac{\delta\Delta}{\Delta_o} = \frac{\delta\lambda}{\lambda_o} = 0; \qquad (13)$$

that is, the field of view is infinite at particular angles of theta. More

importantly, the sign of the angle sensitivity changes between the sectors defined by Equation (12). Along the optic axis and the orthogonal axis the angular sensitivities are, respectively:

$$\left.\frac{\delta\lambda}{\lambda_o}\right|_{\theta=0} = -\frac{\sin^2 i}{2n_o^2} \tag{14}$$

and

$$\left.\frac{\delta\lambda}{\lambda_o}\right|_{\theta = 90^o} = +\frac{\sin^2 i}{2n_o n_e} \tag{15}$$

Along the optic axis the angular sensitivity is the same as a MI as is to be expected. Orthogonal to the axis, however, the difference in curvature between the ordinary and extraordinary waves reverses the angular sensitivity.

The reversal of the sign of the angular sensitivity in orthogonal directions is the physical phenomena that allows construction of a system of much lower angular sensitivity. The WFB element (see Figure 2) is constructed by splitting a normal element in half and rotating the second half so that the optic axes of the two halves are orthogonal. In addition, a half wave plate is inserted between the halves with its axis along the bisector of the angle between the optical axes of the crystals. Since the effect of a half wave plate is to rotate a plane of polarization by twice the angle between the polarization direction and the axis of the wave plate, the effect of the wave plate is to rotate, for states of polarization, the optic axes of the two halves back to parallel. Hence, the on axis time delay of the WFB and a simple element are the same, but for off axis rays the time delay is

$$\Delta = \frac{d}{2c}\left\{n_e\left[1 - \left(\frac{a^2}{n_o^2} + \frac{b^2}{n_e^2}\right)\right]^{1/2} - n_o\left[1 - \left(\frac{a^2+b^2}{n_o^2}\right)\right]^{1/2}\right\}$$
$$+ \frac{d}{2c}\left\{n_e\left[1 - \left(\frac{a'^2}{n_o^2} + \frac{b'^2}{n_e^2}\right)\right]^{1/2} - n_o\left[1 - \left(\frac{a'^2 + b'^2}{n_o^2}\right)\right]^{1/2}\right\} \tag{16}$$

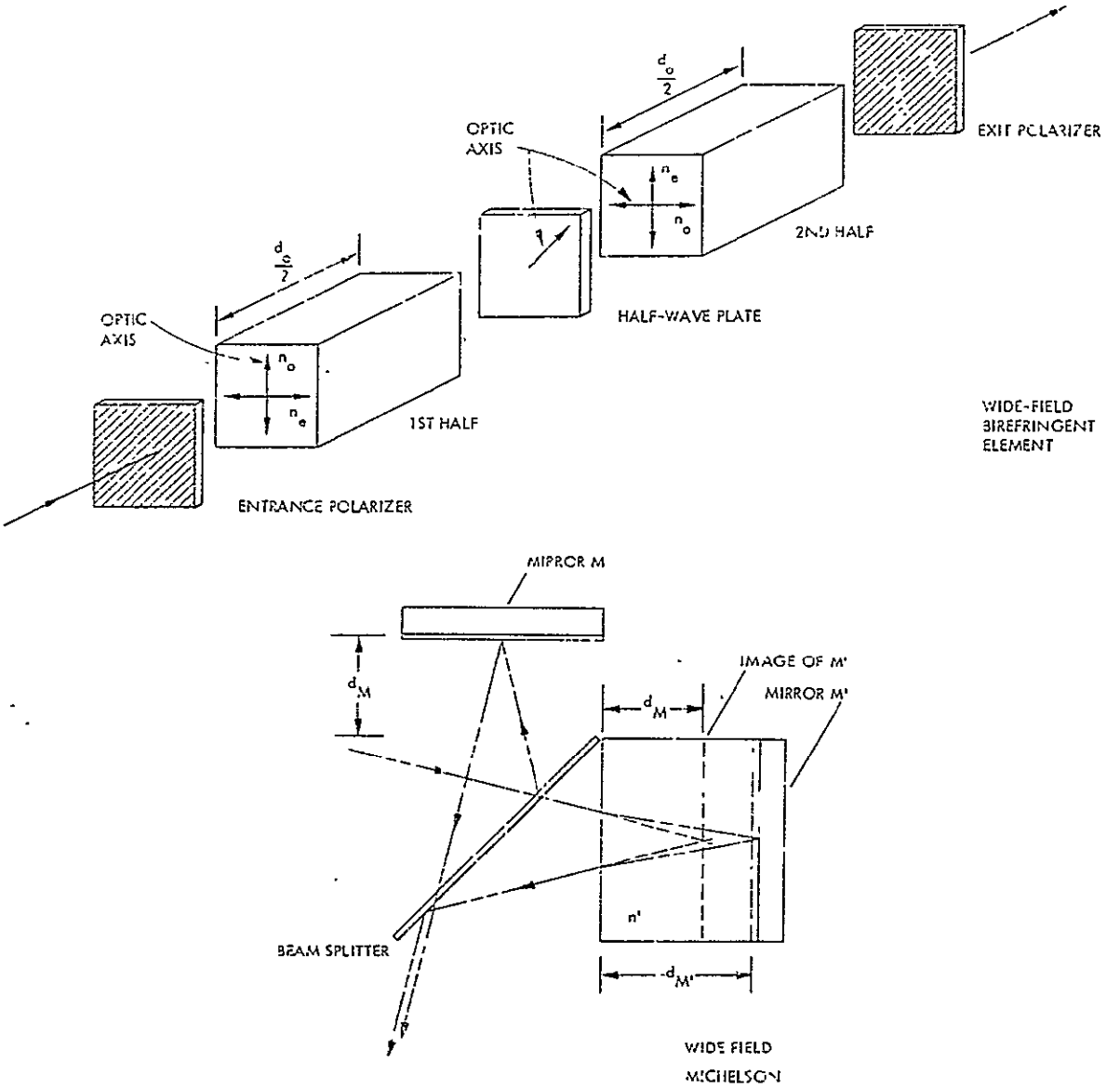


Figure 2. Optical schematic of a wide field birefringent element and a wide field Michelson interferometer.

where

$$a' = \sin i \cos (\theta + 90) \tag{17}$$

and $b' = \sin i \sin (\theta + 90)$.

To the second order

$$\Delta = \Delta_o \left(1 - \frac{1}{4n_o^2} \left(\frac{n_e - n_o}{n_e}\right) \sin^2 i\right) \tag{18}$$

or

$$\frac{\delta\lambda}{\lambda_o} = - \frac{1}{4n_o^2} \left(\frac{n_e - n_o}{n_e}\right) \sin^2 i \ . \tag{19}$$

Since the difference between the ordinary and extraordinary indices is less than one, the wide field construction has the desired result of decreasing the angular sensitivity.

Note that Equation (19) implies that by decreasing the index difference, the angular sensitivity can be indefinitely decreased. This is not the case because of the higher order terms in the expansion. It is, however, worthwhile to consider the WFM, an analogous system whose field effects are determined by fourth order terms. The wide field Michelson, shown in Figure 2, achieves its larger field of view by a block of material placed in front of M', of length, $d_{M'}$, such that the image of M', through the b lock, is optically coincident with M. The time delay[7] is

$$\Delta = \frac{2}{c} \left[n \, d_M \left(1 - \frac{\sin^2 i}{n^2}\right)^{1/2} - n' \, d_{M'} \left(1 - \frac{\sin^2 i}{n'^2}\right)^{1/2} \right]. \tag{20}$$

To the fourth order

$$\Delta = \frac{2}{c}\left[(n\,d_M - n'd_{M'}) - \frac{\sin^2 i}{2}\left(\frac{d_M}{n} - \frac{d_{M'}}{n'}\right) - \frac{\sin^4 i}{8}\left(\frac{d_M}{n^3} - \frac{d_{M'}}{n'^3}\right)\right] . \tag{21}$$

Equation (21) was carried to fourth order because the second order term is zero under the wide field condition which occurs when

$$d_{M'} \simeq \frac{n'}{n}\, d_M, \tag{22}$$

so that

$$\Delta = \Delta_o\ (1 + \frac{\sin^4 i}{8n^2n'^2}), \tag{23}$$

where

$$\Delta_o = \frac{2}{c}(nd_M - n'd_{M'}), \tag{24}$$

and thus

$$\frac{\delta\lambda}{\lambda_o} = \frac{\sin^4 i}{8n^2n'^2} . \tag{25}$$

If the lengths of the WFM are not exactly adjusted, so that

$$d_{M'} = \frac{n'}{n}\, d_M + \varepsilon\,\frac{\Delta_o}{n} , \tag{26}$$

then

$$\frac{\delta\lambda}{\lambda_o} = \frac{\sin^2 i}{2n\ n'}\,\varepsilon + \frac{\sin^4 i}{8n^2n'^2} . \tag{27}$$

Clearly for the case of the WFM the field of view does not become infinite as ε goes to zero, but is limited by the fourth order term. A similar effect should be expected for the WFB. It will be demonstrated below that for a well aligned WFB the angle sensitivity is virtually identical to an appropriately misadjusted WFM.

III. FIELD EFFECTS OF A SIMPLE BIREFRINGENT PLATE

In order to analyze the field effects of not only perfect WFB elements, but also elements which are misaligned or mismatched in length, it is necessary to examine the higher order properties of a simple birefringent plate. The expression for the time delay, equation (10), will be rewritten as:

$$\Delta = \Delta_o \left\{ \frac{1}{q} \left[1 - x^2(1 - (1-p^2)z) \right]^{1/2} - \frac{p}{q} \left[1 - x^2 \right]^{1/2} \right\} \tag{28}$$

in which

$$\Delta_o = \frac{d}{c} (n_e - n_o), \tag{a}$$

$$p = \frac{n_o}{n_e}, \tag{b}$$

$$q = 1-p, \tag{c}$$

$$x = \frac{\sin i}{n_o}, \tag{d}$$

$$z = \sin^2\theta. \tag{e}$$

(29)

Symbolically this is written

$$\Delta = \Delta_o f(x, p, q, z) \tag{30}$$

and has the power series expansion

$$\Delta = \Delta_o \left[1 + c_1 x^2 \, 0(1) + C_2 x^4 0(2) + \ldots + C_n x^{2n} 0(n) + \ldots \right] \tag{31}$$

The coefficients, C_n, are the coefficients of S^n in the expansion of $(1-S)^{1/2}$. Specifically

$$C_n = - \frac{2 \, (2n-2)!}{4^n \, n!(n-1)!} \tag{32}$$

The factors, $0(n)$, are n^{th} degree polynomials in z. From the expansion of the form $(1-S)^{1/2}$ it may be seen that

$$0(n) = \frac{1}{q}\left[(1-(1-p^2)z)^n - p\right] . \tag{33}$$

The first few terms of expansions (32) and (33), are given by:

$$c_1 = -1/2 \qquad 0(1) = 1-(1+p)z \tag{a}$$

$$c_2 = -1/8 \qquad 0(2) = 1-2(1+p)z+q(1+p)^2z^2 \tag{b}$$

$$c_3 = -1/16 \qquad 0(3) = 1-3(1+p)z+3q(1+p)^2z^2\ q^2(1+p)^3z^3 . \tag{c}$$

(34)

Before going further it is worth noting one point, namely that $0(2)$ is nonzero when $0(1)$ is zero. Clearly $0(1)$ is zero when

$$z = \frac{1}{1+p} \tag{35}$$

or when

$$\tan\theta = \pm \left(\frac{n_e}{n_o}\right)^{1/2} . \tag{36}$$

At this point

$$0(2) = -p . \tag{37}$$

Therefore, to fourth order, for $\tan^2\theta = \left(\frac{n_e}{n_o}\right)$,

$$\frac{\delta\lambda}{\lambda_o} = \frac{\sin^4 i}{8n_o^4}\, p . \tag{38}$$

The fourth order term indicates a deviation from the straight line asymptotes at $\tan^2\theta = \pm n_e/n_o$. There are still arbitrarily large angles of incidence for which the time delay sensitivity is zero, but they occur at different thetas.

IV. FIELD EFFECTS OF A WELL ALIGNED WIDE FIELD ELEMENT

The analysis of the single plate may now be used to calculate the properties of a perfect WFB element. Using the formalism of (30), the time delay of a WFB element is

$$\Delta = \frac{\Delta_o}{2}\left[f(x, p, q, z) + f(x, p, q, z')\right] , \tag{39}$$

where, as before,

$$z = \sin^2\theta \tag{40}$$

and

$$z' = \sin^2(\theta+90^o) = \cos^2\theta . \tag{41}$$

As before, expression (39) may be expanded in a power series in the variable x;

$$\Delta = \Delta_o\ 1+C_1 x^2 O_{WF}(1)+C_2 x^4 O_{WF}(2)+\ldots+C_n x^{2n} O_{WF}(2n)+\ldots \tag{42}$$

The coefficients, C_n, are the same as (32), the coefficients of S^n in the expansion of $(1-S)^{1/2}$. The factors, $O_{WF}^{(n)}$, however, are more involved for the wide field element. They are:

$$O_{WF}(n) = \frac{1}{2q}\left[(1-(1-p^2)z)^n + (1-(1-p^2)z')^n - 2p\right] . \tag{43}$$

In this expansion of Δ, the angle of incidence dependency is represented by the power series in x. The theta dependency, the variation of field width as a function of azimuthal position, is included in the polynomial factors, $O_{WF}(n)$. As will be shown, each of these factors has a dominant term which is independent of theta, and smaller terms which are functions of sin 2θ. Specifically this theta independent term will be the sensitivity along the $\theta=0^o$ ray, and is particularly easy to evaluate.

For this special case we have

$$z = \sin^2\Theta = 0$$
$$z' = \cos^2\Theta = 1, \qquad (44)$$

and (43) reduces to

$$O_{WF}(n)\,\Big|_{\Theta=0} = \frac{p^{2n} - 2p + 1}{2q} \cdot \qquad (45)$$

To sixth order in the sine of the angle of incidence, the sensitivity is given as:

$$\frac{\delta\lambda}{\lambda_o} = -\frac{1}{4}\,q\,\frac{\sin^2 i}{n_o^2} - \frac{1}{16}\,(q - p^2\,(1+p))\,\frac{\sin^4 i}{n_o^4} - \frac{1}{32}\,\left(\frac{p^6 - 2p+1}{q}\right)\,\frac{\sin^6 i}{n_o^6} \cdot \qquad (46)$$

Since q is small and p is a number near unity for most crystals, the factors,

$$O_{WF}(n)\,\Big|_{\Theta=0} \simeq \begin{cases} q/2 & n = 1 \\ 1 - n. & n > 1 \end{cases} \qquad (47)$$

Along the $\Theta = 0^o$ ray, therefore, the angle sensitivity is approximately given by

$$\frac{\delta\lambda}{\lambda_o} = -\frac{q}{4}\,\frac{\sin^2 i}{n_o^2} + \frac{1}{8}\,\frac{\sin^4 i}{n_o^4} + \frac{1}{8}\,\frac{\sin^6 i}{n_o^6} \cdot \qquad (48)$$

Equation (48) resembles, at least to fourth order, a slightly misadjusted WFM (equation 27). An interesting consequence of equation (47) is that the coefficients of x^n decay slowly. In particular,

$$C_n\,O_{WF}(n)\,\Big|_{\Theta=0} = \frac{2\,(2n-2)!\,(n-1)}{4^n\,n!\,(n-1)!} \qquad n>1 \qquad (49)$$
$$= \frac{2}{4^n}\,\binom{2n-2}{n} \cdot$$

The first few such factors are

$$C_n\,O_{WF}(n)\,\Big|_{\Theta=0} = \frac{1}{8} = .125 \quad n = 2$$
$$= \frac{1}{8} = .125 \quad n = 3 \qquad (50)$$
$$= \frac{15}{128} = .117 \quad n = 4 \cdot$$

The consequence of this is that for moderate field widths low order expansions will not be accurate.

In order to consider the general form, (43), including the theta dependency it is convenient to expand $O_{WF}(n)$ using the binomial theorem. This results in

$$O_{WF}(n) = 1 + \frac{1}{2q} \sum_{i=1}^{n} (-1)^i (1-p^2)^i \binom{n}{i} (\sin^{2i}\theta + \cos^{2i}\theta) \quad n>1. \tag{51}$$

The factors

$$S_i = \sin^{2i}\theta + \cos^{2i}\theta \tag{52}$$

may be expressed as

$$S_i = 1 - \sum_{k=1}^{i/2} A_i(k) \sin^{2k} 2\theta \tag{53}$$

in which the coefficients, $A_i(k)$, may be determined from the relations:

$$\begin{aligned} A_i(1) &= i/4 && i>2 \\ A_{i+1}(k) &= A_i(k) - \frac{1}{4} A_{i-1}(k-1) . && k>1,\ i>2 \end{aligned} \tag{54}$$

The first few expressions for the sums, S_i, are

$$\begin{aligned} S_1 &= \sin^2\theta + \cos^2\theta = 1 \\ S_2 &= \sin^4\theta + \cos^4\theta = 1 - \frac{1}{2}\sin^2 2\theta \\ S_3 &= \sin^6\theta + \cos^6\theta = 1 - \frac{3}{4}\sin^2 2\theta \\ S_4 &= \sin^8\theta + \cos^8\theta = 1 - \sin^2 2\theta + \frac{1}{8}\sin^4 2\theta \end{aligned} \tag{55}$$

Substituting expressions such as (55) into (51) enables the evaluation of the factors, $O_{WF}(n)$, including the theta dependency. Again the first few such factors are:

$$
\begin{aligned}
O_{WF}(1) &= q/2 \\
O_{WF}(2) &= \frac{p^4 - 2p+1}{2q} - \frac{(1+p)^2 q}{4} \sin^2 2\theta \\
O_{WF}(3) &= \frac{p^6 - 2p+1}{2q} - \frac{3}{8}(1+p)^2 q\,(1+p^2) \sin^2 2\theta \\
O_{WF}(4) &= \frac{p^8 - 2p+1}{2q} - \frac{1}{2}(1+p)^2 q\,(1+p^2+p^4) \sin^2 2\theta + \frac{1}{16}(1+p)^4 q^3 \sin^4 2\theta .
\end{aligned}
\tag{56}
$$

For p very near unity, equations (56) become, respectively:

$$
\begin{aligned}
O_{WF}(1) &= q/2 \\
O_{WF}(2) &\simeq -1 - q \sin^2 2\theta \\
O_{WF}(3) &\simeq -2 - 3q \sin^2 2\theta \\
O_{WF}(4) &\simeq -3 - 6q \sin^2 2\theta + q^3 \sin^4 2\theta .
\end{aligned}
\tag{57}
$$

Each of the factors, $O_{WF}(n)$, may be seen to consist of a constant, theta independent, term and theta dependent terms which are even powers of $\sin 2\theta$. This is the decomposition into the independent and dependent terms, for which the independent terms represent the behavior at $\theta = 0^\circ$. Additionally, they represent the behavior at all multiples of 90° around the element.

Although these theta independent components decrease slowly, as indicated by equations (50), the theta dependent terms decrease rapidly by comparison. This is because the first appearance of a $\sin^{2n} 2\theta$ term is in the $O_{WF}(2n)$ factor. It is

therefore already multiplied by x^{4n}, or $\sin^{4n} i/n_o^{4n}$. In addition, the coefficient of $\sin^{2n} 2\Theta$ in the expression for $O_{WF}(2n)$ is:

$$\frac{(-1)^n (1+p)^{2n} q^{2n-1}}{4^n} . \tag{58}$$

For p close to unity this is approximately

$$(-1)^n q^{2n-1} \tag{59}$$

which decreases very quickly as n increases.

The time delay expression, equation (42), could then be written as

$$\Delta = \Delta_o \left\{ 1 + \Delta (i) \Big|_{\Theta = 0} + \Delta (i, \Theta) \right\} . \tag{60}$$

This implies that the sensitivity is decomposed as $\frac{\delta\lambda}{\lambda_o} = \frac{\delta\lambda}{\lambda_o}\Big|_{\Theta = 0} + \frac{\delta\lambda}{\lambda_o}(\Theta)$. (61)

The first term on the right hand side of (61) has been given, to sixth order, by (46) or approximately by (48). The second term, the theta dependent term, is given approximately by:

$$\frac{\delta\lambda}{\lambda_o} (\Theta) = \frac{q}{8} \frac{\sin^4 i}{n_o^4} \sin^2 2\Theta + \frac{3q}{8} \frac{\sin^6 i}{n_o^6} \sin^2 2\Theta . \tag{62}$$

To fourth order the total sensitivity is then:

$$\frac{\delta\lambda}{\lambda_o} = - \frac{q}{4} \frac{\sin^2 i}{n_o^2} + \frac{1}{8} \frac{\sin^4 i}{n_o^4} (1+q \sin^2 2\Theta) . \tag{63}$$

At this point some insight into the angle sensitivity characteristics of the WFB element can be obtained. Clearly the second order expansion, equation (19), is no longer valid when the second and fourth order terms are equal in magnitude.

This occurs when

$$\sin^2 i_e = 2\,|q|n_o^2 \tag{64}$$

$$i_e = \sin^{-1}\left(2\,|q|n_o^2\right)^{1/2} \tag{65}$$

For this angle the wavelength shift, to second order, would be

$$|\delta\lambda|_{i_e} = \frac{\lambda_o}{2}\,q^2 \,. \tag{66}$$

Since filters or spectrometers are normally used with the maximum wavelength shift equal to the full width at half maximum (FWHM), the second order expansion should be valid for filters with FHWM less than the limit given by equation (66).

For quartz and calcite, the most commonly used birefringent materials, i_e and δ_λ at 5000Å are (9.7°, .08799Å) and (45°, 34.7Å) respectively. The second order approximation is, therefore, almost always valid for wide field calcite elements, but is often not valid for quartz elements. In fact, when a quartz element is used beyond 9.7° even the sign of the wavelength shift predicted by the second order approximation is incorrect.

It should be noted that all the terms in the expansion for Δ (equation 42 et. al.), with the possible exception of the second order term, are positive, Since the coefficients do not decrease rapidly, the overall convergence of the series is, therefore, slow. Detailed calculations of wide field systems must consider terms of higher order than the fourth, especially when the second order term is negative.

V. MISALIGNMENT AND MISMATCH EFFECTS

The variation of the angular sensitivity with azimuth, theta, is negligible for well aligned WFB elements. If the crystals are not perfectly orthogonal, however, then the situation changes markedly and the field of view develops a strong azimuthal dependence. It is worthwhile to investigate these effects of non-orthogonal crystals not only to determine the error sensitivities, but also because angular misalignment offers a mechanism for cancellation of the second and fourth order terms, thereby extending the field of view for some angles.

When the two crystals in the WFB element are non-orthogonal the time delay is again given by equations (39) through (43) but with

$$z' = \sin^2 (\theta+90+\alpha) = \cos^2 (\theta+\alpha) \tag{67}$$

in which α is the misalignment angle. The time delay is again expanded as a power series in x^2, the angle of incidence dependency, but the polynomial factors, $O_{WF}(n)$, need to be modified. In place of equation (51) the expression for $O_{WF}(n)$ now becomes:

$$O_{WF}(n) = 1 + \frac{1}{2q} \sum_{i=1}^{n} (-1)^i (1-p^2)^i \binom{n}{i} (\sin^{2i}\theta + \cos^{2i} (\theta + \alpha)) n>1. \tag{68}$$

By analogy with equation (52) we will define

$$S_i (\alpha) = \sin^{2i}\theta + \cos^{2i} (\theta + \alpha). \tag{69}$$

Letting

$$R = \sin^2\alpha \cos 2\theta + \frac{1}{2} \sin 2\alpha \sin 2\theta, \tag{70}$$

we have

$$\cos^2 (\theta + \alpha) = \cos^2\theta - R \tag{71}a.$$

and

$$\cos^{2i}(\theta + \alpha) = \cos^{2i}\theta + \sum_{k=1}^{i} (-1)^k \binom{n}{k} R^k \cos^{2i-2k}\theta \qquad (71)$$

The factors, $S_i(\alpha)$, may then be expressed as the well aligned component, S_i, plus a correction sum:

$$S_i(\alpha) = S_i + \sum_{k=1}^{i} (-1)^k \binom{n}{k} R^k \cos^{2i-2k}\theta \ . \qquad (72)$$

This means that the polynomial factors, $O_{WF}(n)$, may also be decomposed into a well aligned component plus an alignment error component. The alignment error component has the form:

$$O_{WF}(n;\alpha) = \frac{1}{2q} \sum_{i=1}^{n} (-1)^i (1-p^2)^i \binom{n}{i} \sum_{k=1}^{i} (-1)^k \binom{i}{k} R^k \cos^{2i-2k}\theta \qquad (73)$$

For small misalignments, the factor R may be approximated by

$$R \simeq \alpha \sin 2\theta \ , \qquad (74)$$

and terms in R^2 or higher powers of R may be neglected. Equation (73) is then replaced by:

$$O_{WF}(n;\alpha) \simeq \frac{\alpha \sin^2\theta}{2q} \sum_{i=1}^{n} (-1)^i (1-p^2)^i \binom{n}{i} i, \qquad (75)$$

which equals

$$n\alpha \sin 2\theta \frac{(1+p)}{2} p^{2n-2} \qquad (76)$$

Again letting p be very near unity,

$$O_{WF}(n;\alpha) \simeq n\alpha \sin 2\theta \ . \qquad (77)$$

Adding this term into equations (57) yields, respectively:

$$O_{WF}(1) \simeq \frac{q}{2} + \alpha \sin 2\theta$$

$$O_{WF}(2) \simeq -1 - q \sin^2 2\theta + 2\alpha \sin 2\theta$$

$$O_{WF}(3) \simeq -2 - 3q \sin^2 2\theta + 3\alpha \sin 2\theta \qquad (78)$$

$$O_{WF}(4) \simeq -3 - 6q \sin^2 2\theta + q^3 \sin^4 2\theta + 4\alpha \sin 2\theta .$$

Including the misalignment terms, the angle sensitivity relation is to fourth order:

$$\frac{\delta\lambda}{\lambda_o} \simeq -\frac{1}{4} \frac{\sin^2 i}{n_o^2} (q + 2\alpha \sin 2\theta) + \frac{1}{8} \frac{\sin^4 i}{n_o^4} (1 - 2\alpha \sin 2\theta + q \sin^2 2\theta). \qquad (79)$$

Developed this way it is easy to see the relative magnitudes of the theta independent component, the well aligned theta dependent component, and the misalignment theta dependent component. As may be seen, an angular misalignment can be chosen to cancel the other components for certain directions. This will be discussed further subsequently, but first we will show a similar effect for a length mismatch between the two halves of the WFB element.

Suppose that the lengths of the crystal parts are

$$d_1 = \frac{d_o}{2} + Td_o \qquad (80)$$
$$d_2 = \frac{d_o}{2} - Td_o .$$

The time delay is then

$$\Delta = \frac{\Delta_o}{2} \left[f(x, p, q, z) + f(x, p, q, z') \right] + T\Delta_o \left[f(x, p, q, z) - f(x, p, q, z') \right] . \qquad (81)$$

The first term is the familiar wide field time delay, while the second term is

similar except for the minus sign. Again, this second term may be expanded as a power series in x^2, but with the polynomial factors:

$$O_{WF}(n;T) = \frac{1}{2q}\left[(1 - (1-p^2)\ z)^n - (1-(1-p^2)\ z')^n\right]. \tag{82}$$

If we consider only a length mismatch this becomes

$$O_{WF}(n;T) = \frac{1}{2q}\sum_{i=1}^{n} (-1)^i\ (1-p^2)^i\ \binom{n}{i}\ (\sin^{2i}\theta - \cos^{2i}\theta). \tag{83}$$

The factors

$$D_i = \sin^{2i}\theta - \cos^{2i}\theta \tag{84}$$

may be expressed as

$$D_i = -\cos 2\theta\left[1 + \sum_{k=1}^{\frac{i-1}{2}} B_i\ (k)\ \sin^{2k}2\theta\right] \tag{85}$$

in which the coefficients, $B_i\ (k)$, may be determined from the relations:

$$B_i\ (1) = \frac{2 - i}{4} \qquad i>2 \tag{86}$$

$$B_i\ (k) = \frac{1}{4}\ B_{i-2}(k-1) - A_{i-1}(k) \qquad i>3,\ k>1$$

The terms, $A_{i-1}(k)$, are as defined by equations (54).

Since we are assuming that the length mismatch, T, is small and that p is near unity, we may neglect all but the first term of (83):

$$O_{WF}(n;T) \approx n \cos 2\theta\ . \tag{87}$$

The contribution of the length mismatch to the time delay expansion is, therefore,

$$\Delta\ (T) = \Delta_o T \cos 2\theta\ \{c_1 x^2 + 2c_2 x^4 + \ldots + nc_n x^{2n} + \ldots\} \tag{88}$$

Combining this with the fourth order expansion, equation (63), the total sensitivity, including mismatch, is:

$$\frac{\delta\lambda}{\lambda_o} = -\frac{1}{4}\frac{\sin^2 i}{n_o^2}\ (q + 2T\cos 2\theta) + \frac{1}{8}\frac{\sin^4 i}{n_o^4}\ (1 + q\sin^2 2\theta - 2T\cos^2\theta). \tag{89}$$

Equation (89) shows that the effect of a length mismatch is very similar to a rotation misalignment except that the field of view expands in different angular ranges.

We will not consider the combined effect of rotation misalignment and length mismatch since the additional terms would be on the order of (αT), which can be safely neglected.

VI. THROUGHPUT ADVANTAGES OF WIDE FIELD SYSTEMS

With the relations between relative wavelength shift and the angle of incidence and azimuth derived in the sections above it is possible to demonstrate the relation of the WFB element to the F-P, MI, and WFM. For negative crystals there are two regimes as indicated by equations (64) through (66), depending upon whether the second or fourth order terms dominate the angle sensitivity.

When the second order term dominates

$$\frac{\delta\lambda}{\lambda} < q^2/2 \ , \tag{90}$$

and the maximum incident angle for an allowed relative wavelength tolerance is obtained from

$$\sin i = \frac{2n_o}{|q|^{1/2}} \left(\frac{\delta\lambda}{\lambda}\right)^{1/2} . \tag{91}$$

Relation (9) is a similar result for the F-P and MI. So that the ratio of acceptance angles of a WFB to a F-P or MI is:

$$I_2 = \left(\frac{2n_o}{n_{FP}}\right) \frac{\sqrt{2}}{|q|^{1/2}} . \tag{92}$$

Since q is always less than one, the WFB has a larger field than the F-P and MI. The gain in throughput for equal area devices, neglecting vignetting is

$$G_2 = I_2^{\,2} = \frac{2}{q} \left(\frac{2n_o}{n_{FP}}\right)^2 . \tag{93}$$

For a calcite filter ($q = .12$) the gain in throughput is approximately 46 for $n_o = 1.67$ and $n_{FP} = 1$.

Regardless of the sign of q, when the fourth order term dominates the wavelength sensitivity the maximum incident angle for the WFB is obtained from

$$\sin i = 2\, n_o \left(\frac{1}{2}\frac{\delta\lambda}{\lambda}\right)^{1/4} = n_o \left(8\,\frac{\delta\lambda}{\lambda}\right)^{1/4} \tag{94}$$

and the ratio of acceptance angles with respect to a F-P or M is

$$I_4 = \left(\frac{n_o}{n_{FP}}\right) \left(2\frac{\lambda}{\delta\lambda}\right)^{1/4} , \qquad (95)$$

and the gain in throughput is

$$G_4 = \left(\frac{n_o}{n_{FP}}\right)^2 \left(2\frac{\lambda}{\delta\lambda}\right)^{1/2} . \qquad (96)$$

Since $\frac{\delta\lambda}{\lambda}$ can be very small the WFB has the potential of significant throughput gain with respect to the F-P and MI. Comparison of relations (94) and (25) shows that the angular field and hence throughput of the WFB and WFM are equivalent when the fourth order term dominates.

For positive crystals another possibility exists because the wavelength sensitivity is zero at the point the second and fourth order terms balance in magnitude, thus the wavelength sensitivity has an extremal which is determined by

$$\frac{\partial}{\partial} (\sin i) \left(\frac{\delta\lambda}{\lambda}\right) = 0 , \qquad (97)$$

and occurs when

$$\sin i = n_o\sqrt{q} . \qquad (98)$$

From relations (98) and (63) the extremal value is

$$\frac{\delta\lambda}{\lambda} = - q^2/8 . \qquad (99)$$

From the equality condition (equation 64) the relation between angular field and

wavelength tolerance in the optimized condition is

$$\sin i = 2 n_o \left(2 \frac{\delta\lambda}{\lambda}\right)^{1/4} \tag{100}$$

So that the gains in field and throughput compared to a F-P or MI are

$$I_o = 2 \left(\frac{n_o}{n_{FP}}\right) \left(\frac{1}{2} \frac{\lambda}{\delta\lambda}\right)^{1/4}, \tag{101}$$

and

$$G_o = 4 \left(\frac{n_o}{n_{FP}}\right)^2 \left(\frac{1}{2} \frac{\lambda}{\delta\lambda}\right)^{1/2}, \tag{102}$$

respectively. The optimized condition thus yields a factor of 2 gain in throughput over the fourth order condition. It should also be clear that a WFM can be adjusted to increase its throughput by an appropriate adjustment.

The angular characteristics of the WFB can be illustrated by a pair of log-log plots as shown in Figure 3. The upper plot, (3a), shows sine i versus relative wavelength, $\delta\lambda/\lambda_o$, for quartz and calcite. The true values (solid lines) are compared with the 2nd and 4th order approximations (dashed lines). In addition, a curve for a crystal of optimal q (equation 100) is also presented for comparison. On the right hand scale of Figure 3a are values of the angle of incidence and on the top scale are the values of filter bandpass at λ_o = 5000Å for telecentric systems whose maximum shift equals the wavelength tolerance. Immediately below, in Figure 3b, are the corresponding plots of the ratio of incidence angle (left scale) and the throughput ratio (right scale) versus the relative wavelength tolerance.

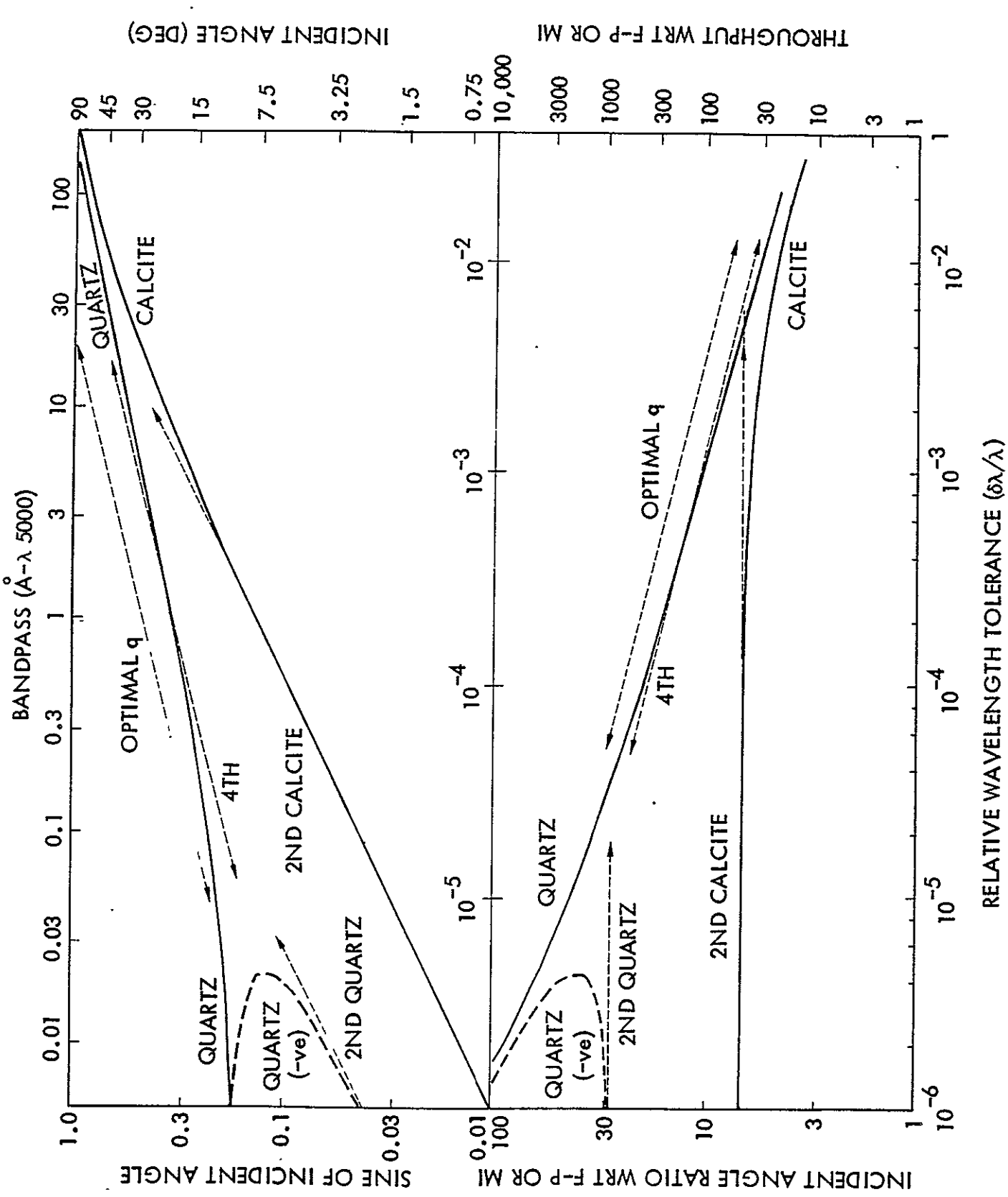


Figure 3. Angle of incidence and throughput advantage versus wavelength tolerance and bandpass for quartz, calcite and optimal q wide field elements.

Figure 3 clearly indicates the regions for which the 2nd and 4th order approximations are valid for quartz and calcite. It also presents the size of the gains in acceptance angle and throughput which are possible with a WFB with respect to a F-P or MI. A calcite WFB can be built for the range of relative wavelength tolerance from 10^{-5} to 10^{-3} where a throughput gain of nearly 50 can be realized. For quartz, the wavelength tolerance range from 2×10^{-4} to 6×10^{-3} is practical, for which throughput gains from 40 to 300 are possible.

The arguments above for optimimization of q are not strictly academic. Although the selection of natural crystals is strictly limited it should be possible to create sheet plastic materials with desired q. Uniaxial sheet plastic wide field elements can be built by lamination.

Even if it is not possible to select q it is clear from equations (79) and (89) that the factor multiplying the second order term can be made zero by appropriate rotation or mismatching of crystal lengths, respectively.

Adjustment of the angle or lengths has the greatest effect when the second order term dominates the wavelength sensitivity. To estimate the field gain consider equation (79) and suppose α is adjusted such that at 45° and 225° the second order term vanishes. Then the wavelength sensitivity function will have an oval shape with major and minor axes given by

$$\sin i_{major} = 2\, n_o \left(\frac{1}{2}\frac{\delta\lambda}{\lambda_o}\right)^{1/4} \tag{103}$$

and

$$\sin i_{minor} = \sqrt{2\, n_o}\, \left(\frac{\delta\lambda}{\lambda}\right)^{1/2} \tag{104}$$

The misadjustment results in an increase in throughout

$$W = \frac{|q|}{2}^{1/2} \quad (\frac{\lambda o}{2\delta\lambda})^{1/4} \tag{105}$$

compared to the properly aligned element. For a calcite element of 0.1Å bandpass (λ5000Å) this gain is about 3, which means that the field in one direction increases by a factor of 6. Using the higher order terms yields somewhat higher estimates.

VII. FRINGE PATTERNS OF WIDE FIELD SYSTEMS

An important use of the wavelength sensitivity calculations performed above is for the explanation of the fringe patterns produced by wide field elements of different materials in various states of alignment or length mismatch. The patterns are easily produced by observing an element over a diffuse monochromatic source. Since a bright fringe occurs whenever the optical path difference is an even multiple of halfwaves, the fringe pattern is a contour map of the wavelength variation. That is, the radius vector to a given bright fringe is determined by the angle of incidence and azimuth required for a particular wavelength shift. It also should be noted that unless the illuminating source has two wavelengths the contours are unsigned; the fringes can be caused by either a positive or negative wavelength shift. Of course, with two wavelengths reversal of the wavelength shift will cause a reversal of the fringe colors

Once the fringe patterns are understood the observer has a wonderfully sensitive tool to use for aligning and matching crystals. Further, the patterns are very beautiful and are in our opinion worth understanding on purely esthetic grounds.

The bright fringes occur from the fourth order expansion whenever

$$m = R\left[-\frac{\sin^2 i}{4n_o^2}\,(q + 2\alpha \sin 2\theta + 2T \cos 2\theta) + \frac{\sin^4 i}{8n_o^4}\right], \qquad (106)$$

where m is a signed integer, and

$$R = \frac{\Delta_o c}{\lambda_o} . \qquad (107)$$

For equation (106) it has been assumed that the effects of q, α, and T on the 4th order term in sine i can be neglected. In order to calculate the fringe pattern

accurately, higher order terms must be considered in regions where the wavelength sensitivity is slowly varying. Contours on gently slopes can move a great deal with very small changes in inclination of the slope. For this reason a computer program called PAT has been written to solve an eighth degree polynomial equation. Since all terms higher than second order in sine i are positive (equation (46), the general shape of the fringe pattern can usually be determined from equation (106), which means most of the properties of the fringe patterns follow from the characteristics of a simple equation in x^2

$$x^4 - 2q'x^2 - \frac{8m}{R} = 0, \tag{108}$$

where

$$q' = q + 2\alpha \sin 2\Theta + 2T \cos 2\Theta . \tag{109}$$

Fringes arise from solutions of relation (108) for all possible integer values (both positive and negative) of m for fixed q, α, and T. The solutions are limited, however, by the condition that x^2 be real, positive and less than n_o^{-2}. The limitations on x^2 are because sine i must be less than one. As is well known, there are two solutions of a quadratic equation which involve, respectively, a positive and negative contribution from a surd. For equation (108) the surd is

$$Q = \left[(2q')^2 + 32m/R\right]^{1/2} . \tag{110}$$

Because x^2 must be real Q must also be real. This means that when q' is small, that is, when $q'^2 < 8m/R$. (111)

only solutions for positive m are possible, while if q' is large

$$\text{i.e., } q'^2 > \frac{8m}{R} \tag{112}$$

both positive and negative solutions are possible. The condition that x^2 be positive places further limits on m depending on whether q is positive or negative. The valid solutions for positive and negative m with the assumption that either (111) or (112) holds are summarized in Table I.

For perfect elements (α and T equal zero) of small q, Table I shows that there are only solutions for positive m regardless of the sign of q. Thus the fringe pattern for small q is generated by

$$\sin i = n_o \left(\frac{8m}{R}\right)^{1/4} . \qquad (113)$$

Note that the circular fringes have radii proportional to the fourth root of the integers and, therefore, form a pattern similar to that of a WFM.

For perfect elements of large negative relative birefringence, Table I shows that there exist only solutions for positive m and that the fringe pattern is generated by

$$\sin i = 2n_o \left(\frac{m}{R|q|}\right)^{1/2} . \qquad (114)$$

The fringe radii of this pattern are proportional to the square roots of the integers, and thus the pattern is similar to that of a F-P or MI.

For the case of large positive values of q, Table I suggests that there are three possible sets of fringe patterns. For negative m there is the expected set of fringes given by

$$\sin i = 2n_o \left(\frac{|m|}{Rq}\right)^{1/2} , \qquad (115)$$

		$\lvert q' \rvert < \left(\frac{8m}{R}\right)^{1/2}$		$\lvert q' \rvert > \left(\frac{8m}{R}\right)^{1/2}$	
SGN m	SGN SURD	$q' > o$	$q' < o$	$q' > o$	$q' < o$
+	+	$4\sqrt{2\left(\frac{m}{R}\right)^{1/2} + q'}$	$4\sqrt{2\left(\frac{m}{R}\right)^{1/2} - q'}$	$2\,q'\left(1 - \frac{2m}{Rq'^2}\right)^{-1}$	$4\,\lvert m\rvert\left(1 + \frac{2m}{Rq'^2}\right)^{-1}$
+	−	NO SOL'N	NO SOL'N	NO SOL'N	NO SOL'N
−	+	NO SOL'N	NO SOL'N	$4\,\frac{\lvert m\rvert}{Rq'}\left(1 + 2\,\frac{\lvert m\rvert}{Rq'^2}\right)^{-1}$	NO SOL'N
−	−	NO SOL'N	NO SOL'N	$2\,q'\left(1 + \frac{2m}{Rq'^2}\right)^{-1}$	NO SOL'N

TABLE I - SOLUTIONS FOR X^2

and also the set given by

$$\sin i = n_o \left[2q \left(1 + \frac{2|m|}{Rq2}\right) \right]^{1/2} . \qquad (116)$$

Further, there are positive m fringes given by

$$\sin i = n_o \left[2q \left(1 - \frac{2|m|}{Rq2}\right) \right]^{1/2} . \qquad (117)$$

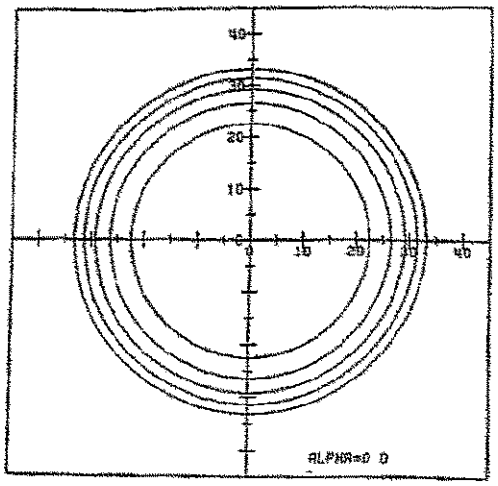
The fringes indicated by expressions (116) and (117) occur because of the cancellation of the second and fourth order terms, and the angle

$$i = \sin^{-1} (2qn_o^2)^{1/2} . \qquad (118)$$

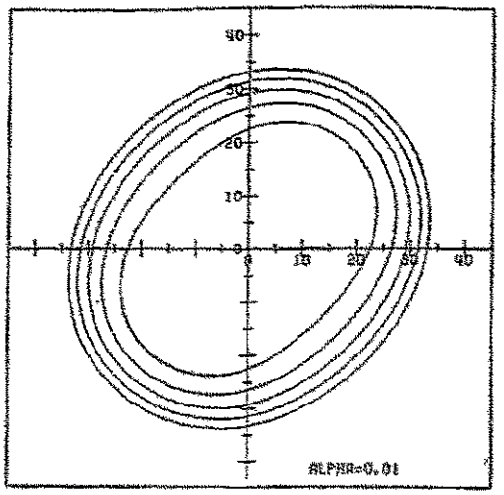
represents the transition between second and fourth order fringe patterns.

Unfortunately, positive crystals with high relative birefringence are not normally made into wide field elements, so that these fringes have not been observed, at least by our group. However, it is straightforward to calculate the appearance of any fringe pattern using PAT. Shown in Figures 4a, 5a, and 6a are the locations of the first few bright fringes of a 1Å (λ 5000, $\frac{\delta\lambda}{\lambda_o} = 2 \times 10^{-4}$) wide field element of quartz (small q), calcite (large negative q) and reverse calcite, (large positive q), respectively. In all PAT plots negative fringes are dashed. Reverse calcite is imaginary material with reversed ordinary and extraordinary indices from calcite.

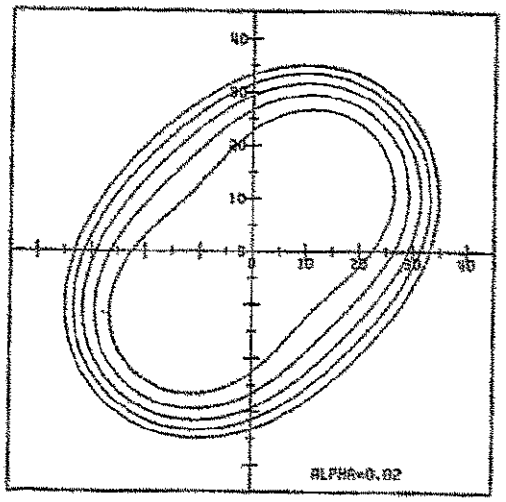
For elements that are misaligned, α is non zero and the value q' oscillates in magnitude when 2α is less than $|q|$ and oscillates both in magnitude and sign when 2α is greater than $|q|$. First consider the case of a large negative q material. When 2α is less than $|q|$, $|q'|$ will decrease in the 45^o-225^o direction and increase



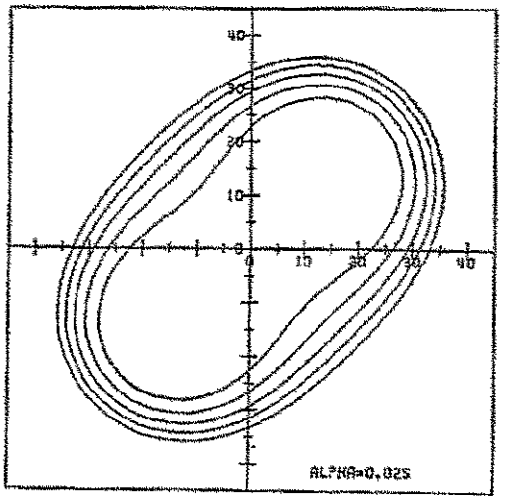
4(A)



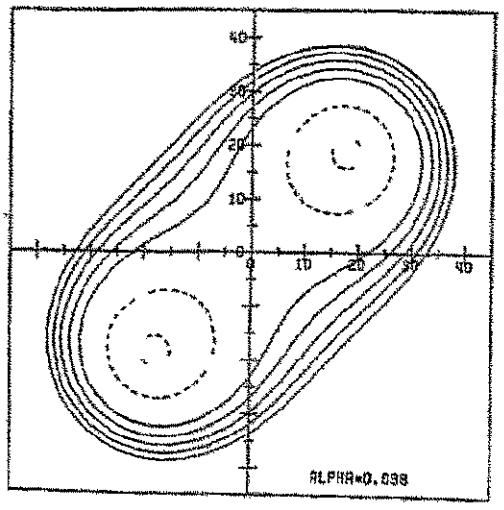
4(B)



4(C)

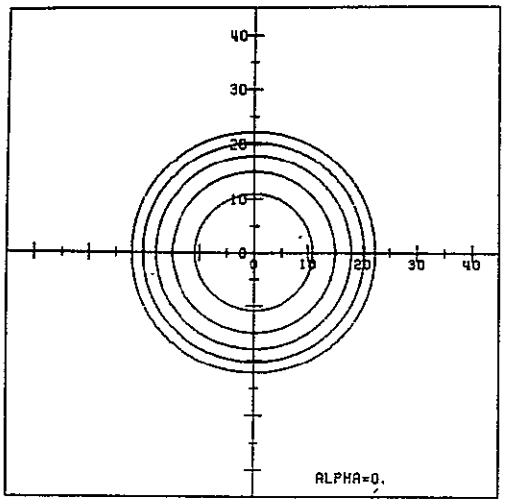


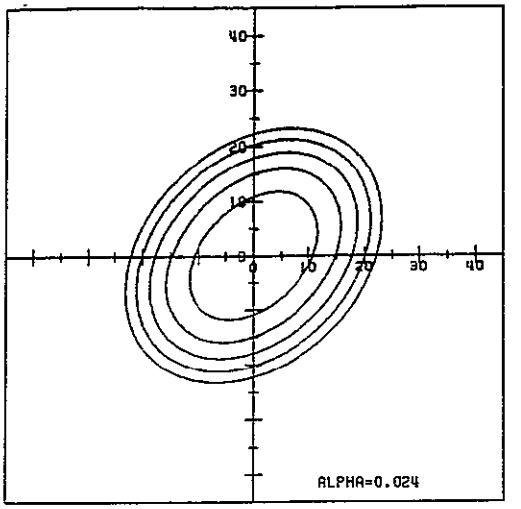
4(D)

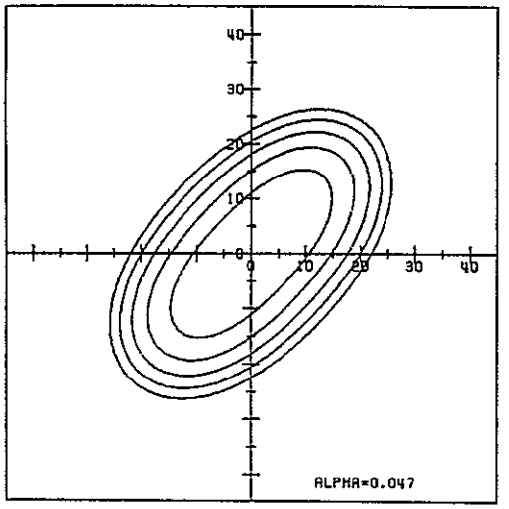


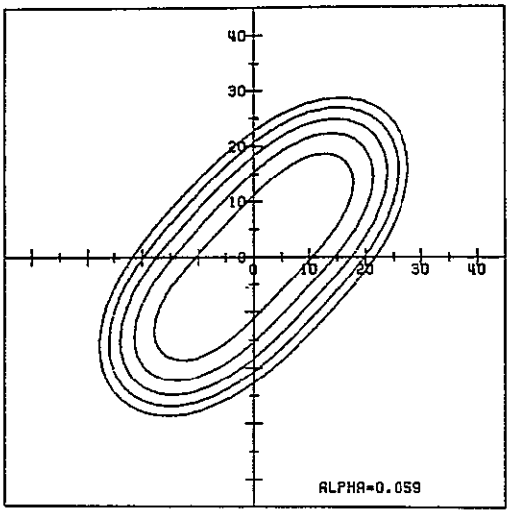
4(E)

Figure 4. Fringe patterns for misalignments of 0, $.4\alpha_o$, $.8\alpha_o$, α_o, and $1.5\alpha_o$, respectively, of a quartz element.









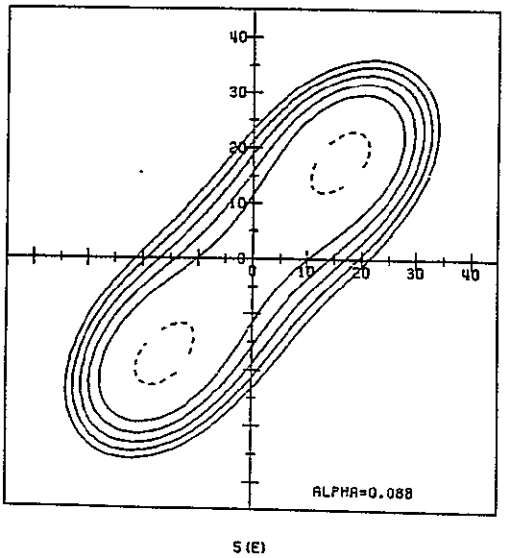
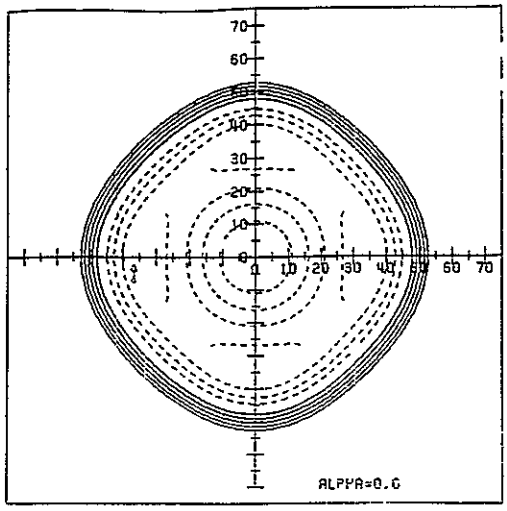
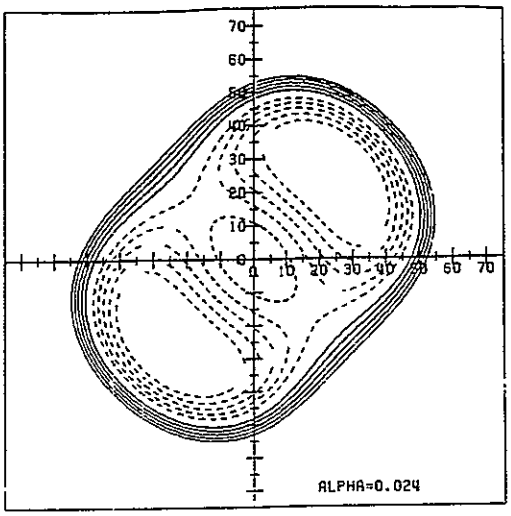


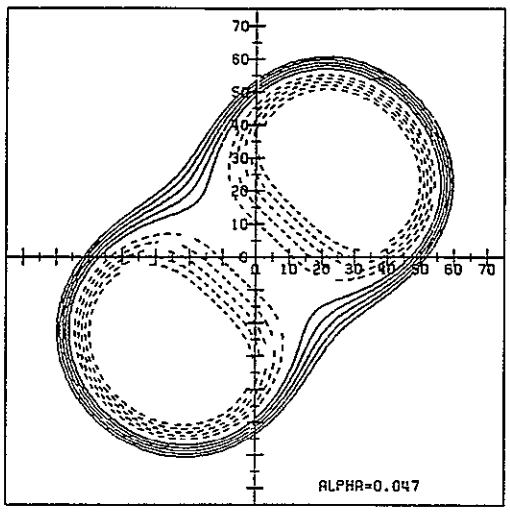
Figure 5. Fringe patterns for misalignments of 0, .2q, .4q, .5q, .75q, respectively, of a calcite element.



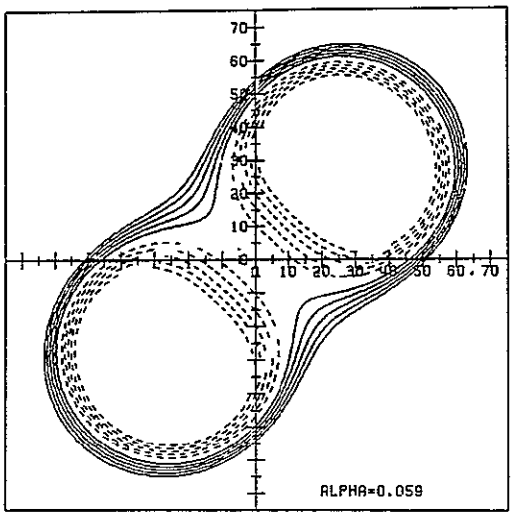
6 (A)



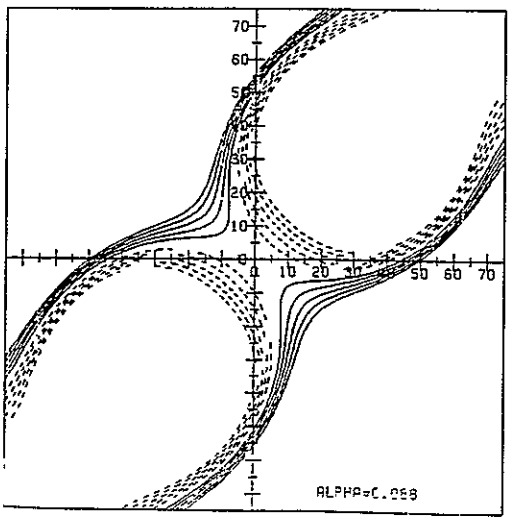
6 (B)



6 (C)



6 (D)



6 (E)

Figure 6. Fringe patterns for misalignments of 0, .2q, .4q, .5q, .75q, respectively, of a reverse calcite element.

in the 135^{o}-315^{o} direction. Since the fringe radii are inversely proportional to $|q'|^{1/2}$, the fringes expand along the 45^{o}-225^{o} direction and contract along the orthogonal direction. The expansion and contraction are, therefore, along and orthogonal to the polarizer direction. When α is positive the expansion is along the 45^{o}-225^{o} direction, while when α is negative it is along the 135^{o}-315^{o} line. Note also that as α approaches $|q|$ the fringes do not extend to infinity, but are limited by the small q solution which is almost independent of q and depends mostly on m and R.

When 2α is greater than $|q|$, Table I, shows that fringes of negative m become possible. The negative m solutions give rise to egg-shaped fringes that nest inside the extended positive fringe system. The fringe pattern for a 1Å calcite (large negative q) WFB element is illustrated in Figures 5b, c, d, e where α equals .2q, .4q, .5q, .75q, respectively. The pattern of Figure 5e is very similar to that of a biaxial crystal.

For misaligned WFB elements of large positive q the situation is somewhat different. When 2α is less than q the inner fringes behave in a similar manner to those of a negative q material except the elongation is along the 135^{o}-315^{o} line and the contraction along the 45^{o}-215^{o} line for a positive rotation. Both the positive and negative m outer fringes expand along 45^{o}-225^{o} direction and contract along in the 135^{o}-315^{o} direction because (see Table I) they are proportional to q' rather than inversely proportional to $|q'|^{1/2}$ as are the inner fringes. When 2α is greater than q the angular range of the negative m fringe system is limited because there are no solutions for both negative m and q'. What occurs is illustrated in Figures 6b, c, d, e for a 1Å reverse calcite (large positive q) WFB element. The

α values are the same as for Figure 5.

When q is small the fourth order term in sine i dominates and quite a different process dominates the rotational misalignment fringe patterns. This can be seen using equation (79) and the approximation that q can be neglected with respect to α. We then have:

$$m/R = (1 - 2\alpha \sin 2\theta)\, x^4 - 2\alpha \sin 2\theta\, x^2 . \qquad (119)$$

When α is small compared to 1/R only positive m solutions are possible and the fringes will elongate along the line 45°-225° and contract along 135°-315°. If α is increased further a negative fringe system will occur inside the elongated positive fringes. The result for small q is that regardless of sign the fringes will elongate along the 45°-225° or 135°-315° line depending upon whether α is positive or negative. The first negative fringe can be expected when α is approximately equal to:

$$\alpha_o = {}^1/R . \qquad (120)$$

The value α_o in (120) is determined by the extremal value of equation (110) for the case m equals minus one. Shown in Figures 4b, c, d, e are the fringe patterns of a quartz 1Å WFB element for α equal .4 α_o, .8 α_o, α_o, and 1.5 α_o.

The patterns of Figures 4, 5, and 6 represent a reasonably complete catalog of the misalignment fringe patterns of uniaxial WFB elements. When WFB elements are mismatched in length a very similar set of patterns arise as should be expected from comparing equations (79) and (89). The elongations are along the 0°-180° or

90^{o}-270^{o} lines or along or orthogonal to the optic axes depending upon whether T is positive or negative.

In general a WFB element if it is removed from a birefringent filter or in an early state of manufacture will have both angular and length misadjustment. Therefore, the field of view may be elongated in an arbitrary direction. By careful adjustment of the element halves with respect to each other, however, it should be possible to get either circular fringes or more likely fringes elongated along or orthogonal to the optic axis. From Figures 4, 5 or 6 and equation (89), the elongations can be estimated. Nearly perfect fringes will occur for large (small) q WFB elements when α and T are less than 1/20 q $(1/10\ R^{-1/2})$. For calcite and quartz this means the tolerance on angle and length match of a $1\mathring{A}$ are $(.32^{o}$, .004 cm) and $(.114^{o}$, .026 cm), respectively.

VIII. DISCUSSION

In this paper the main intent has been to explain the very significant throughput advantages of WFB filters and the possible modifications of the field of view for sources that are extended in one direction. The paper is not intended to be a complete discussion of off axis error sources and light loss in wide field elements as was paper III[8] of this series for on-axis effects. The effects of imperfect waveplates, waveplate field, and crystals with optic axis not in the plane of the crystal face will be considered in a further paper of this series.

This paper has resulted in large part from experience gained in constructing three wide field of view birefringent filters. The filters are a pure Lyot 10Å bandpass quartz filter, a redundant alternate partial polarizer (RAPP) 60Å bandpass polyvinyl alcohol filter, and a modified alternate partial polarizer filter (APP) with 60mÅ bandpass. All of these filters use the plastic waveplates discussed in paper I through III[9,10] and the internal apodization of pulse structure concepts introduced in paper IV[11]. However, a new general technique for partial polarizer filter design was used for the RAPP and modified APP designs which also will be described in a later paper of this series.

The quartz and polyvinyl alcohol filters are wide field and tunable from 4500-7000A and have diameters of 70m and 35mm, respectively. They exhibit the large throughput advantages predicted above. The polyvinyl filter is especially interesting because its elements are constructed by laminations of one wave sheets of stretched polyvinyl purchased from Polaroid Corporation. Such lamination techniques should allow experimentation with optimally stretched sheet plastic materials.

IX ACKNOWLEDGEMENT

It is a pleasure to acknowledge Harry Ramsey for building the wide field elements and demonstrating the beautiful fringe patterns of various perfect and imperfect wide field elements; Ralph Reeves for the design and construction of the filters; and Steve Schoolman and Thomas Pope for their activities in developing thermal compensation systems. Finally, this work has been supported by the NASA cometary and solar programs under contracts NASW-3107 and NAS8-32805; the NSF Astronomy Program under Grant AST77-21084; and Lockheed Independent Research funds.

REFERENCES

1. B. Lyot, Comptes Rendus 197, 1593 (1933)
2. B. Lyot, Ann. Astrophys. 7, No. 1-2, 31 (1944)
3. J. W. Evans, J. Opt. Soc. Am. 39, 229 (1949)
4. A. M. Title, T. P. Pope, M. E. Ramsey, S. A. Schoolman, NASA Report CR 156674, (1976).
5. R. L. Hilliard and G. G. Shepherd, J. Opt. Soc. Am., 56, 362 (1966)
6. M. Francon and S. Mallick, Polarization Interferometers, Wiley - Interscience, 1971
7. H. H. Zwick and G. G. Shepherd, Appl. Optics, 10, 2569 (1971)
8. A. M. Title, Appl. Optics, 14, 445, (1975).
9. A. M. Title, Solar Physics, 33, 521 (1973).
10. A. M. Title, Appl. Optics, 14, 229, (1975).
11. A. M. Title, Appl. Optics, 15, 2871 (1976).

NASA CR 156674

DEVELOPMENT OF BIREFRINGENT FILTERS FOR SPACEFLIGHT

NAS Contract No. NAS 5-22368

JULY 1976

Prepared By
LOCKHEED PALO ALTO RESEARCH LABORATORIES
3251 HANOVER STREET
PALO ALTO, CALIFORNIA 94304

Dr. Alan M. Title Harry E. Ramsey
Dr. Thomas P. Pope Steven A. Schoolman

This study was performed under the auspices at the Solar Physics Quick Reaction and Special Purposes Facility Definition Team (Contract NAS 5-22368) for the Shuttle/Spacelab Payloads Project, Goddard Space Flight Center, NASA.

Prepared For
GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND 20771

TABLE OF CONTENTS

GENERAL INTRODUCTION

This report is on work done for the NASA Solar Physics Spacelab Quick Reaction and Special Purpose Facility Definition Team on the critical problems of a space qualified birefringent filter suitable for use on an Hα/magnetograph. This report uses data, fabrication techniques and computer programs for the analysis of birefringent systems previously developed by the Lockheed Solar Observatory. The construction design philosophy of the filter modules tested as well as the construction techniques were based on previous experience with numerous birefringent filters.

The critical problem for flight of a birefringent filter is the shock mounting of the calcite. The design presented here bonds the calcite block with silicon rubbers to the calcite holder. The calcite together with its all necessary polarizers and rotating achromatic plates are mounted together in units called a filter module. By using a set of modules containing calcite crystals of differing lengths, a filter can be produced. A description of the modules is contained in Section I. Also described in Section I is a container for the filter modules, which can be used both to hermetically seal the system or contain an index matching oil.

A second serious problem with birefringent filters has been wavelength drifts introduced by thermal variations. The traditional approach to the thermal drift problem has been a hardwired temperature controller and oven enclosure for the calcite. Section II reports on the response of a filter element while being controlled by the Lockheed Temperature Control.

A tunable filter such as proposed for the Hα/magnetograph offers the possibility of using the tuning elements to compensate for any thermal drifts thus eliminating any need for a hardware temperature controller if the rate of change of wavelength with temperature is known. Section III describes the determination of the wavelength sensitivity to temperature of calcite. The

Lockheed Universal filter has been built with temperature sensors on the calcite elements. Section IV describes operation of the filter using a software control algorithm instead of a hardware temperature controller.

Section V briefly reviews some radiation considerations of filter systems. Also included is an appendix that indicates future areas of research for filter materials and systems.

I. SPACE HARDENED FILTER MODULES

INTRODUCTION

The goal of the space hardening experiments was to contain calcite elements in a cell that would sustain both shock testing and thermal cycling. It was desired that the cell also contain the necessary components for a complete tunable module. A preliminary test model was used to verify the cell design, a prototype was designed to verify the tunable module, and an oil sealed prototype to verify a complete filter assembly.

In addition, optical cementing techniques were developed to allow reliable cement bonds between calcite to calcite, calcite to glass, and calcite to polyvinyl alcohol (PVA). Bonding was of some concern because previous calcite cements had a tendency to cause fracture when bonded assemblies were thermally cycled. Reliable cements allow construction of filters that do not require index matching oils.

SHAKE TEST

Preliminary Test Unit

Before making a full 30 mm aperture 1Å filter module, a test unit with 25 mm aperture was constructed. The test unit consisted of 4Å and 2Å wide field elements. Table I lists the components of the Mark I test unit. The test unit was constructed with and sealed in its housing with GE RTV 602. The test units were shaken to the specification given in Astronomy Spacelab Payload Accommodations and Interfaces Handbook (May 1975), Figure 4.2.2.2-1 (reproduced in Figure 1). The unit was shaken in three orthogonal axis, where the y axis was parallel to the filter optical axis. A sample shake test chart is shown in Figure 2.

The first test unit shake, 2-25-76, showed separation of the waveplates and their cover glasses. The separation was evident from fringe patterns

seen when the filter was mounted over a diffuse helium source. A photograph of the fringes is shown in Figure 3. The filter was disassembled and the separation of the waveplates was verified. The failure was attributed to the poor bond between RTV 602 and the polyvinyl waveplate material.

Although the separation occurred, the calcite showed no evidence of damage nor did any of the components rotate with respect to each other. The spectral performance of the Mark I version of the test unit did not change.

The test unit was rebuilt using Eastman Kodak HEF-4 optical cement to bond the waveplate material. The glass covers of the polyvinyl in intermediate positions in the filter were eliminated. The optical component list is shown in Table II. The filter was bonded in the cell with RTV 602 as before. The mechanical drawing of the Mark II system is shown in Figure 4. A photograph of the cell and shake fixture are shown in Figure 5.

Kodak HEF-4 is a non-hardening optical cement. Calcite elements have been cemented with HEF-4 and temperature cycled between 20° and 60°C.

On 3-17-76 the unit was reshaken. No separation was observed. No evidence of any adverse effects was shown by the filter. Spectral performance was not affected.

Prototype Filter Module

After the successful experience with the test unit a flight prototype unit was constructed. The flight prototype unit has an achromatic half waveplate and a rotating polarizer. Its optical components are listed in Table III. A mechanical diagram and a photograph are shown in Figures 6 and 7.

On 5-4-75 the prototype unit was successfully shaken. The calcite did not rotate and the spectral performance was unchanged.

The calcite used was fabricated by Halle Nacht and has two small cracks on its edge. This calcite was especially chosen to see if the cracks would increase. They did not.

TABLE I

MARK I COMPONENTS LIST

Material	Thickness (cm)	Orientation	Comment
1. Glass	.203		Bonded to 2- RTV 602
2. 1/4 λ (PVA)	.00254	45°	3
3. Glass	.203		4
4. Calcite	.254	0°	5
5. Glass	.203		6
6. 1/2 λ (PVA)	.00254	45°	7
7. Glass	.203		8
8. Calcite	.254	90°	9
9. Glass	.203		10
10. HN-38 (PVA)	.00254	0°	11
11. Glass	.203		12
12. Calcite	.127	90°	13
13. Glass	.203		14
14. λ/2 (PVA)	.00254	45°	15
15. Glass	.203		16
16. Calcite	.127	0°	17
17. Glass	.203		18
18. λ/4 (PVA)	.00254	45°	19
19. Glass	.203		

TABLE II

MARK II COMPONENTS LIST

Material	Thickness (cm)	Orientation	Comment
1. Glass	.203		Bonded to 2 - HEF4
2. 1/4 λ (PVA)	.00254	45°	3
3. Calcite	.254	0°	4
4. 1/2 λ (PVA)	.00254	45°	5
5. Calcite	.254	90°	6
6. HN-38 (PVA)	.254	0°	7
7. Calcite	.127	90°	8
8. λ/2 (PVA)	.00254	45°	9
9. Calcite	.127	0°	10
10. λ/4 (PVA)	.00254	45°	11
11. Glass	.203		

TABLE III

Prototype Components List

Material	Thickness (cm)	Orientation	Comment
1. Glass	.203	Rotates	Bonded to 2 with HEF-4
2. HN-38	.00254	as	3
3. Glass	.203	unit	
4. Space	.114		
5. Glass	.203		Bonded to 7 with HEF-4
6. $\lambda/4$ (PVA)	.00254	45°	8
7. Calcite	.5588	0°	9
8. $\lambda/2$ (PVA)	.00254	15°	10
9. $\lambda/2$ (PVA)	.00254	75°	11
10. $\lambda/2$ (PVA)	.00254	15°	12
11. Calcite	.5588	90°	13
12. HN-38	.00254	0°	14
13. Glass	.203		

SEALED FILTER ENCLOSURE

After the successful shake test of the prototype filter module the spur gears were replaced by bevel gears which were more suitable for the rotary feed-through system used for the sealed enclosure assembly. The enclosure was designed as a hermetic or oil sealed case for the birefringent filter. The enclosure actually built is just sufficiently long to contain the prototype filter module, but a larger enclosure with more feed-throughs could contain any number of filter modules.

The enclosure was built with O- ring seals at all the non-rotating joints. The enclosure components and the complete enclosure are shown in Figs. 8 and 9. For a flight unit epoxy or welded seals might be judged more suitable. The rotary feed-through necessary to control the tuning module has no rotary oil seal. Rather, rotary motion is transferred through an intermediate plate that rocks. The plate is welded to a metal bellows which in turn is welded to the side of the feed-through so that there is no piercing connection between the input and output shafts. No magnets of any kind were used in the seal.

The enclosure was tested for gas tightness by flooding the exterior with helium while pumping on the interior with a vacuum pump. A helium leak detector in the output part of the pump showed no leakage signal. The tuning motor was operated as the unit was pumped. Following the gas seal tests the enclosure was filled with index matching oil and placed in a vacuum chamber. After forty hours in the chamber at 10^{-6} torr no oil was detectable on the surfaces of the container. The motor was operated periodically while the unit was in the vacuum chamber.

On 6-25-76 the oil filled enclosure was shaken to the shuttle shake profile. At the time of the shake test the enclosure contained the prototype filter module. No oil leakage was observed after the shake. The unit was then replaced in the vacuum chamber and left in vacuum for 16 hours. The motor was periodically rotated. After removal from the vacuum chamber, no leakage was observed.

The oil inspection was carried out by wiping a clean cotton swab over all metal parts. The exposed optical surfaces were visually inspected. Any oil filter would create interference colors or tint, none were observed.

A-R COATING POSSIBILITIES

A major problem with ground based filters has been leaking and contamination of the index matching fluid or oil in the filter. In space, the consequences of leakage are much more severe. One solution to the leakage problem is to completely seal the filter as described above. Another is to eliminate the index matching fluid.

The index matching fluid serves two purposes. The first is to eliminate fresnel reflection losses at the interfaces. The second is to eliminate image displacement caused by rotating tilted optical elements in the filter. During this study a cementing technique has been developed which allows reliable stress-free assembly of calcite to calcite and calcite to glass. The index of the cement is 1.57. The index of the cement used for the waveplate is 1.54.

If a Lyot filter is constructed from a set of turnable modules of the form half wave, quarter wave, calcite, half wave, calcite, polarizer where the first half wave plate rotates and all the other components are cemented together, there are only four air-glass interfaces per module. A plot of maximum transmission, neglecting polarizer losses, versus number of modules is shown in Fig. 10 for several values of per surface reflection loss. Shown in Fig. 11 is the per surface reflection loss for a sample four layer antireflection coating. From 4300Å to 6600Å the per surface reflection loss is less than 0.2%. This spectral region spans the useful region of currently available polarizing materials. Recent developments suggest coatings may be available that have reflectance less than 0.2% over the entire range from .3 to 2.2 microns.

For the filter design suggested above, only the half wave plates rotate. If the half wave plates are 0.25 cm thick and are mounted in 5-cm ring bearings with less than 1.25×10^{-3} cm axial play, then the displacement of the optical beam through the plate is

$$d = t\theta \left(\frac{n-1}{n}\right)$$

$$d = (2.5 \times 10^{-1}) \left(\frac{1.25 \times 10^{-3}}{2.5}\right) \left(\frac{.52}{1.52}\right)$$

$$d = 4.14 \times 10^{-5} \text{ cm}$$

An image scale of 40 arc seconds per cm is such that 0.1 arc second occupies 2.5×10^{-3} cm, the size of a typical diode array element. A 40 π cm scale thus is nearly 2 orders of magnitude greater than any image motion caused by rotation.

To verify that a cemented filter would operate well, one of the observatory Halle Hα filters was disassembled, degreased, and reassembled with cemented components and A-R windows. The filter is now in routine use in the observatory patrol system.

The rebuilt Halle has been repeatedly cycled between 10°C and 50°C. At present it is left outside with the power off during the evening hours and is also uncovered during normal operation in order to subject the cement bonds to larger than normal thermal loads.

TEMPERATURE CONTROLLER TESTS

PURPOSE

The purpose of the thermal controller tests was to evaluate the ability of a stand alone temperature controller to maintain the temperature of calcite filter elements in the presence of environmental thermal loads and solar heating. The temperature controller used was the Lockheed Filter Controller designed by R. Lindgren. This device senses ambient temperature and supplies a steady state heat input based on the difference between ambient and the control temperature. In addition, it applies a fine control heat input based on a sensor near the calcite crystal. The power input follows the law:

$$HI = k_1(t_c - t_a) + k_2(t_c - t_s)^2 \, \mathrm{sgn}(t_c - t_s) + k_3$$

where k_1, k_2 and k_3 are adjustable parameters, and t_c, t_s and t_a are respective the control, sense, and ambient temperature values. If the sensed temperature is greater than the desired control temperature, the ambient control heat input is reduced. The control input is limited to be less than 20 percent of the ambient heat input.

METHOD OF MEASUREMENT

Two fundamentally different systems were used to measure the characteristics of the controller. One, using a quartz crystal thermometer, measured the temperature in the oven close to the calcite. The second measured the mean temperature through the crystal using a laser probe. The laser probe allowed measurements of possible temperature inhomogenities in the calcite itself. Measurements were made on a .08Å wide field filter element which was installed in a Halle oven. A quartz thermometer probe was mounted in the aluminum block enclosing the calcite (see Figure 12). The temperature was read out on a Hewlett-Packard (Dymec) Quartz Thermometer. The quartz thermometer is repeatable over the short term to ± .0005°C and over times of days to ± .001°C. It is linear over the range used to ± .0008 degrees per degree.

Shown in Figure 13 is an optical schematic of the test system. A 2 mm He-Ne collimated laser beam is sent through the filter and the output detected by a photometer. In addition to the laser beam, a solar beam can be simultaneously shone through the filter. The filter is tunable by rotating a half waveplate behind the entrance polarizer. Both the entrance and exit polarizers are mounted separately from the Halle oven.

As the temperature changes, the index of refraction difference and the length of the calcite changes, so that the laser line is alternately transmitted and extinguished. For this crystal there are approximately 2.22 transmission cycles/^{o}C. Fractions of a cycle can be measured by rotating the half waveplate to an angle where a transmission minima occurs. The locations of transmission minima can be reliably located to one degree in angle, thus the equivalent temperature measurement accuracy is $\pm$ 1/90 of 1/2.22 or $\pm$ .005^{o}C.

CHARACTERISTICS OF THE CONTROLLER

Four identical controllers were evaluated. Controllers were set to control at approximately 30^{o}C, 36^{o}C, 42^{o}C and 48^{o}C. In a normal room environment, all the controllers would maintain an oven temperature of $\pm$ .001^{o}C as indicated by the quartz thermometer readout. When the temperature reading of the quartz thermometer was stable, the laser transmission was also stable. The effects of thermal environment of the temperature controller itself was tested by separately placing the calcite oven and the temperature control electronics in a thermal chamber that could be cycled from 0^{o} to 50^{o}C. The temperature of the calcite oven changed by -.0028^{o}C/^{o}C of ambient over the ambient range from 0^{o}C to 50^{o}C. The electronics caused a change of .001^{o}C/^{o}C ambient. With precision components the temperature control system could be improved by a factor of at least two.

SOLAR EFFECTS ON OVEN TEMPERATURE

With the oven interior temperature well stabilized and the waveplate adjusted so that a transmission minima occurred for the laser beam, an f/17 solar beam was allowed to pass through the filter. The beam was prefiltered by a 8Å three period blocker and an IR rejection filter. There was some tendency for the transmission minima to move a degree in angle or for the calcite to change by .005°C in temperature. However, after repeating the test a number of times, no statistical significance could be obtained for a shift of a degree.

To further test for evidence of solar heating, the 8Å blocker was replaced by a 150Å blocker which increased the flux by a factor of nearly 20. Again, there was some evidence of a change in angle of maximum extinction by a degree in angle. Hence, temperature change, if it exists, must be attributed to a minor IR leakage of the trim filter. In any case, the calcite changed by less than .005°C or .00181Å which is certainly not of any consequence.

TRANSIENT EFFECTS AND THERMAL GRADIENTS

By tracking the transmission of the laser beam as a function of time, it was possible to measure the temperature of the calcite as a function of time. To measure the temperature response on heating the oven was set at a control temperature using a temperature controller. The waveplate was rotated to obtain a transmission minima, and then a new temperature controller was attached to the oven. (Separate temperature controllers were used because each individual controller is tuned to control optimally at a specific temperature. The controller can be fine tuned over a one degree range.) During the heating period, the output of the laser was continuously recorded on a strip chart recorder. As the calcite heated, the laser transmission was modulated, a new minima occurring when the total phase change of the calcite changed π radians (one fringe).

When marks on the recorder chart were made at specific temperatures indicated by the quartz thermometer during the warm-up, a comparison of the thermal response of the filter oven and the calcite crystal was obtained.

A plot of oven temperature and calcite temperature versus time during a transition from 30° to 36°C is shown in Figure 14. As can be seen from the curve, there is almost a degree centigrade difference between the aluminum block and the center of the calcite during the heat-up phase. (The laser beam probed the calcite at its center.) For the run shown, the calcite was a press fit into the aluminum housing.

If the observed temperature difference was occurring across the calcite, about 2 fringes would have been observed along the aperture radius. Under illumination by a full aperture laser beam, the radial phase change would manifest itself by two concentric bright rings which continuously expanded out from the center of the crystal as the crystal heated. When illuminated by a full aperture collimated beam, no ring structure was observed during heat up. To check more quantitatively, the 2mm laser beam monitor was moved to pass through the calcite 1.5 mm from the edge of the crystal and then thermal cycle experiment was repeated. The response of the calcite at the edge was identical to the response at the center of the crystal. Thus, no thermal gradient exists across the calcite even under the fastest heat input condition of the controller, .4°C/min.

To further check the location of the thermal gradient, the element was removed and thermal contact silicon grease was installed in the space between the crystal, its containing tube, and the oven block. The thermal cycle experiment was repeated. The grease reduced the gradient by a factor of two. The rest of the gradient most probably is across the aluminum tape that is wrapped around the calcite.

From the tests on the sample element it can be inferred that a shock mounted and space qualified calcite element would be isothermal, but that a temperature gradient would exist across elastomer mounting substance. The size of the gradient would be dependent upon the thermal input rate. In any case, such gradients can be measured and calibrated.

TEMPERATURE SENSITIVITY OF CALCITE

INTRODUCTION

The sensitivity of the operating wavelength of a birefingent filter to temperature is a critical parameter in the design of an active filter controller which corrects for temperature changes by tuning elements by rotation of polarizers or wave-plates. A knowledge of the temperature sensitivity is also important for the design of a constant temperature enclosure. Somewhat surprisingly, the temperature sensitivity of calcite has not been measured with high accuracy or as a function of wavelength. The only published value being that of Lyot at the Sodium D line.

The thermal effect can be measured by two techniques. The first directly measures the change in wavelength of a transmission peak of a Lyot element due to a change in temperature. The second measures transmission modulation as a function of temperature at a fixed wavelength. The wavelength sensitivity is obtained from the transmission modulation through the relation

$$\frac{\partial \lambda}{\partial T} = \lambda^2 \frac{2\pi}{\varphi} \frac{\partial \varphi}{\partial T} \left[\frac{\lambda}{\left(1 - \frac{\lambda}{\mu} \frac{\partial \mu}{\partial \lambda}\right)} \right] \tag{1}$$

where

$$\varphi = 2\pi \frac{\mu d}{\lambda} \tag{2}$$

and where μ is the index difference, d is the thickness of the crystal, and λ the wavelength. The change in phase, φ, with temperature comes from both change in index and crystal length. That is,

$$\frac{\partial \varphi}{\partial T} = \varphi \left(\frac{1}{\mu} \frac{\partial \mu}{\partial T} + \frac{1}{d} \frac{\partial d}{\partial T} \right) \tag{3}$$

Comparison of equations (1) and (2) shows that the temperature sensitivity is independent of crystal length.

Measurements of both the thermal coefficient of expansion and of index are given in the International Critical Tables; however, they are not sufficiently accurate to determine the wavelength sensitivity. The index values have been fitted by a polynomial by Beckers which is accurate to a part in 10^5 - the accuracy of the index values. The temperature derivative term in the polynomial is on the order of 1×10^{-5} per degree, so that it is not surprising that the temperature sensitivity predicted by the critical values are not accurate to better than 10-20 percent.

To measure the shift in wavelength directly requires a very well calibrated high resolution spectrograph. Therefore, these experiments were carried out at Kitt Peak. To measure the modulation with temperature requires only stable wavelength sources and an accurate photometer. The modulation experiments were performed at Lockheed.

MEASUREMENT TECHNIQUES

In order to measure values of $\partial\lambda/\partial t$, a Halle filter oven was modified to contain a single calcite crystal. The crystal was a normal wide-field Lyot element with a separation between transmission maxima of approximately 1Å. The oven temperature was monitored by the quartz thermometer. The temperature was controlled by a set of four Lockheed temperature controllers with nominal control temperatures of 30, 36, 40, and 48°C.

The filter was mounted vertically over the entrance slit of the main spectograph of the McMath Solar Telescope at Kitt Peak National Observatory (which is a 13.7 meter Czerny-Turner with a spectral resolution 300,000). The filter was fed with an f/60 solar beam. Measurements were made by using the spectrograph in a spectrometer mode scanning over approximately 10Å centered at 4702, 5150, 5850 and 6439Å. A total of 12 scans were co-added for each measurement. At each of the wavelengths, a solar spectrum was taken before and after the measurements at the four temperatures.

The measurements at Lockheed were made using the system used for the thermal controller experiments, where the laser could be replaced by a line source

with a collimator and appropriate isolation filter (Figure 13). The line source measurements were carried out at λ7032, 5461, 4800, 4678, and 4358. Laser measurements were made at λ6328. These measurements used the 0.08Å calcite element - the element used in the thermal controller experiments.

KITT PEAK NATIONAL OBSERVATORY DATA REDUCTION

All of the raw Kitt Peak data were manipulated using a spectral data handling program called REDUCER. The spectral data were corrected for the grating equation, fourier transformed, filtered, and retransformed. The transform operations served to optimally smooth the data. The smoothed data were then analyzed to find all of the locations of intensity maxima and minima. The wavelength differences between corresponding minima or maxima on records made at different temperatures yielded $\partial\lambda/\partial T$.

The measured values for the temperature sensitivity did not show any temperature dependence. The results of the KPNO measurements are shown in Table IV. The least squares straight line fit through the data points is

$$\frac{\partial\lambda}{\partial T} = 7.35 \times 10^{-5} \lambda - .10273 \qquad (4)$$

where $\frac{\partial\lambda}{\partial T}$ is in Å/°C and λ is in angstroms. The mean temperature is 39°C.

LOCKHEED DATA REDUCTION

The experiment at Lockheed directly measured the number of intensity cycles between two temperatures. Because the number of cycles could be obtained with high accuracy it was noted in the first measurements that the fringe count/°C was temperature dependent. Since even Lyot had stated that the temperature sensitivity

Table IV

Temperature Sensitivity of Calcite From KPNO Data

	$\frac{\partial \lambda}{\partial T}$ (39°)
4702	.2433
5150	.2771
5866	.3306
6439	.3704

of calcite was independent of temperature from 18°C to 40°C, numerous measurements were made at each temperature in order to obtain good statistics.

Shown in Table V are the mean temperature and mean sensitivity as well as the least square fit straight line parameters for each of the measured wavelengths. Also shown in Table V are the total number of measurements at each wavelength. Using all the straight line fits of the fringe count per degree at the various temperatures, a mean temperature dependent straight line can be obtained for the fringe count.

$$F/^{\circ}C = \alpha_{\lambda} \ (1 + 5.88 \times 10^{-8} \lambda T) \tag{5}$$

The parameter α_{λ} is excellently fit by

$$\alpha_{\lambda} = 1.494 \times 10^{4}/\lambda - .16443$$

Thus the number of fringes/°C is

$$F/^{\circ}C = (-.16443 + 1.494 \times 10^{4}/\lambda) \ (1 + 5.88 \times 10^{-8} \lambda T) \tag{6}$$

where λ is in angstroms and T in °C. The values given in equation (6) agree with the experimental values to within 0.1% from λ4358 to λ7032. The length of the calcite element used in the experiment was 13.010 cm, so that the number of fringes/cm/°C is

$$F/cm/^{\circ}C = (-.01264 + 1.1483 \times 10^{3}/\lambda)$$

$$\times (1 + 5.88 \times 10^{-8} \ T \lambda) \tag{7}$$

The fringe sensitivity times the free spectral range (peak to peak separation) is just the wavelength sensitivity to temperature. The free spectral range is

$$FSR = \frac{\lambda^{2}}{\mu d \ (1 - \frac{\partial \mu}{\partial T} \ \frac{\lambda}{\mu})} \tag{8}$$

Table V

Fringe Count Parameters and Temperature Sensitivity
of Calcite From Lockheed Experiment

	F/°C	FSR (39°)	$\frac{\partial \lambda}{\partial T}$ (39°)	α_0	$\alpha_1 \times 10^4$	N
4358	3.31	.06575	.2176			8
4678	3.0646	.07887	.2405	3.034	7.76	24
4800	2.96	.08417	.2492			8
5461	2.602	.11524	.2999	2.568	8.712	35
6328	2.225	.1622	.361	2.197	7.114	60
7032	1.994	.20540	.4096	1.958	9.58	24

Note: 1) $F/^\circ C = \alpha_o + \alpha_1 T$

2) N is the number of measurements

The FSR for a centimeter length calculated from the Beckers formula can be fitted to better than a tenth of a percent by

$$FSR = (-2.722 \times 10^{-1} - 1.5065 \times 10^{-6} \lambda + 5.9725 \times 10^{-8} \lambda^2) \quad (9)$$

The product of equations (7) and (9) represents an analytic expression for the temperature sensitivity with an accuracy of a tenth of a percent.

Thus, the final analytic approximation for the temperature sensitivity is

$$\frac{\partial \lambda}{\partial T} = (-3.126 \times 10^2/\lambda + 1.711 \times 10^{-3} + 6.86 \times 10^{-5} \lambda - 7.55 \times 10^{-10} \lambda^2) \times (1. + 5.88 \times 10^{-8} \lambda T) \quad (10)$$

The temperature sensitivity and the number of fringes/cm/°C versus wavelength at 40° are shown in Fig.15. The FSR for a cm of calcite versus wavelength is shown in Fig.16.

Although it is not readily apparent from casual comparison of equations (4) and (10), the KPNO and Lockheed results do agree at 39°C for the region 4700 to 6500Å. Both above and below the KPNO range (4700-6500) the temperature sensitivity departs as can be seen on Figure 15.

AUTOMATIC WAVELENGTH STABILIZATION

INTRODUCTION

The traditional source of wavelength drift in birefringent filters has been inadequate thermal control of the calcite. The temperature controller described above is adequate for control to better than .005° C for periods of weeks to months, which is more than adequate for ground based uses. However, for a universal filter with individual tuning controls on each element, automatic temperature compensation by tuning is possible. Automatic tuning has the advantage of eliminating the temperature controller electronics. Absolute wavelength calibration is also possible by reference to a laser.

LOCKHEED UNIVERSAL TECHNIQUES

The Lockheed Alternate Partial Polarizer Universal Filter (LAPPU) is the first filter system designed to operate without a temperature controller. Instead, temperature sensors are mounted adjacent to the calcite crystals and the elements are tuned to maintain wavelength based on the signals from the sensors and temperature sensitivity of calcite. (Some care has been taken in the design to make the filter respond slowly and uniformly to temperature perturbations.) Although additional software complexity is introduced to compensate for the lack of a stand-alone temperature controller, the software required is a relatively small addition to wavelength referencing and automatic filter assembly which are features which are useful if not necessary for a Universal filter.

For the purpose of auto wavelength referencing, a Helium-Neon laser beam is shown through the filter. The filter control computer can read the values of the laser intensity transmitted by the filter. This data is used by a control program which successively

rotates each element and estimates the position of maximum transmission. Each tuning element is set to its individual position of maximum transmission. Several iterations of the program are sufficient to optimally assemble the filter on λ6328. Periodic use of the laser reference and auto-assembly guarantees that the filter is operating properly and is on wavelength.

In order to operate the filter at any other wavelength than λ6328 either the laser is referenced or the current estimated setting of λ6328 is used to set the locations of the waveplates. Since the temperature sensitivity is known to a part in 1000, the setting capability is better than a degree in angle for the widest element. Narrower elements have corresponding higher resolution. One degree accuracy corresponds to a wavelength setting accuracy of the element FSR/180.

Operationally, to go to a standard wavelength in the spectrum a table of angle offsets from λ6328 is referenced. These values are corrected by the temperature state of the filter. That is

$$\theta_k^i = \theta_k^i (T_0) + \frac{180}{FSR_{6328}^i} \frac{\partial\lambda}{\partial T_{6328}} - \frac{180}{FSR_k^i} \frac{\partial\lambda}{\partial T_k} (T-T_0) \qquad (1)$$

where $\theta_k^i(T_0)$ is the angular offset from λ6328 of element i operating at wavelength k_i at the reference temperature T_0, θ_k^i is the angular offset to which element i is driven to operate at wavelength k, and FSR_k^i is the peak to peak separation of element i at wavelength k. Any nearby wavelength can be obtained by rotating the components proportional to the length of the elements tuned. To maintain wavelength the angular positions are updated when a change in temperature is sensed. The angular change is

$$\Delta\theta_k^i = \frac{180}{FSR_k^i} \frac{\partial\lambda}{\partial T_k} \Delta T \qquad (2)$$

At present analytic approximations for FSR_k^i and $\frac{\partial\lambda}{\partial T_k}$ exist. Analytic expressions for the now tabulated values of $\theta_k^i(T_0)$ will be developed. They can be easily obtained by feeding the filter with a solar beam on the spectrograph bench and using a

version of the wavelength reference auto assembly program that obtains intensity data from the spectrometer output. The only requirement is that the spectrograph is set in the continuum.

MEASUREMENTS

Using the laser reference system built into the Lockheed Universal filter each element was successively spun and the position of maximum transmission measured. After each spin operation, the particular filter tuning element was set to the position of maximum transmission. This operation was carried out continously for 6 hours. All spin data was stored. This included the set position of all elements, the transmission at the set point, and the transmission as a function of angle for each element. After each spin sequence, the filter was set to the position of λ6328 predicted by the temperature compensation program. A comparison of the predicted position and the spin set positions yields the relative error of the temperature compensation routine.

During the testing, all stepping motors were always on, even if not stepping, and the filter was enclosed by a box. In a six hour period the interior temperature of the calcite increased from 20°C to 43°C. During the entire period the difference between the setting of the filter based on the thermal compensation program and the measured setting on the laser line was dominated by the least count errors of the thermal sensors. Ignoring least count effects the filter stayed on wavelength to within .01Å for the entire period.

The test described above is much more severe than ever would be encountered in normal operation. If the filter is not enclosed in a box, the interior temperature does not rise significantly. Normal operation would not operate the motors on much more than a five percent duty cycle.

RADIATION EFFECTS

The radiation problem for birefringent filters comes from two sources. One is the incident solar optical radiation and the other is the penetrating particle and X-ray flux.

Because of the sensitivity of the required blocking filters to solar radiation below 4000Å and the leakage of polarizers above 7500Å, the entrance windows must limit the solar optical flux to the 4000-7500Å band. The filters used by the Skylab H-alpha telescope accomplish this and should be used in front of the birefringent filter. The solar flux behind the Skylab type filter is just the same as normally used on ground-based telescopes.

Penetrating radiation will have to traverse at least 3 cm of aluminum when incident from the side of the filter housing. Radiation along the axis of the filter will have all the optics and the secondary mirror structure in the path which will be greater than the equivalent of 3 cm of aluminum. Measurements of the radiation dose indicates that in a 330 nautical mile polar orbit the dosage is well described by

$$\text{Dose} = 27.4 10^{-t/.19} \text{ rads/day}$$

where t is in cm of aluminum. Radiation dose in a shuttle orbit would be less than this. This relation was obtained by E. Gaines (LMSC) based on data from Trapped Particle Environment of SMM, GSFC X-601-76-97. For a thickness of 3 cm of Al, the dose is

$$\text{Dose (3 cm)} = 3 \times 10^{-15} \text{ rad/day.}$$

Radiation damage occurs in polymer, crystals, and glasses at a radiation dose of 10^5 to 10^7 rads (Space Materials Handbook, Addison-Wesley, 1965).

From the above numbers the likelihood of damage to any of the materials inside the housing is very small.

SUMMARY AND CONCLUSIONS

The basic direction of this study has been guided by the desire of the Quick Reaction and Special Purpose Facility Definition Team for feasibility data on a birefringent filter system for an Hα/magnetograph telescope. The primary requirements were for a 1/6Å or narrower filter that was tunable from Hα to a suitable magnetically sensitive line or lines in the visible. The two principal unresolved problem areas for a space qualified filter were proper shock mounting for the calcite filter elements, and wavelength stability. This study has developed a prototype tunable filter module which has survived the shuttle shake specification.

Experience has shown that the primary cause for wavelength instability is temperature induced wavelength shifts. Tests of the Lockheed temperature controller demonstrate sufficient temperature stability for a facility magnetograph. Tests of the Lockheed Universal filter using thermal sensors on the calcite elements and a software program to control the tuning components have shown that it is possible to temperature compensate without a hardware temperature controller. An advantage of the software system is the ability to incorporate a wavelength reference He-Ne laser to periodically verify the state of the filter. The complete software system for wavelength tuning, thermal control, and wavelength reference can easily be contained by a small sophisticated desk calculator such as an HP 9825. Further studies of the Lockheed Universal filter have shown that with a knowledge of the wavelength dependence of temperature it is possible to actively tune a filter to any desired wavelength with sufficient accuracy ($>$ .01Å) without a hardwired temperature control.

A secondary concern of this study has been the possibility of leakage of index matching oil from the filter. An oil sealed container with a novel rotary feed-through has been constructed and environmentally tested and demonstrated not to leak.

Oil filled versus non-oil filled filter designs have been evaluated. Optical cementing techniques have been developed which allow the majority of

the filter components to be mounted in groups, thus minimizing the number of air-glass interfaces. Calculation of cemented-A-R coated filter designs have shown that it is practical to eliminate the oil in a birefringent filter. A production grease-filled Halle Hα filter has been disassembled and reassembled with cemented components and A-R coated surfaces, but without oil. The A-R-cemented filter performs satisfactorily. Recently announced A-R coatings should allow operation from .35 to 2 microns without oil.

The modular approach to filter design and construction illustrated by the prototype filter module and the Lockheed Universal filter allows the construction of filters of any full width at half-maximum (FWHM) and free spectral range (FSR). Limitations on the full width at half-maximum are imposed by length of calcite. However, the existing Lockheed Universal has a FWHM a factor of three narrower than that desired by the FDT. The FSR is limited only by the number of modules used. The state of the art in blocking filters requires six to eight modules for a magnetograph system.

As a result of the work performed in this study, in the opinion of the authors, a tunable birefringent filter suitable for flight is now within the state of the art. However, A-R coated filters are limited to the range of the A-R coating.

APPENDIX I

The purpose of this appendix is to record my (A.M. Title) judgement on material and techniques for producing birefringent elements.

1. Calcite

At present there seems to be at least four sources of calcite. The calcite comes from mines in Mexico, South Africa, India, and perhaps the USA. We have contacted a half dozen agents and have obtained prices in the range of $750 to $2000 per kilogram. The calcite in the Lockheed Universal, as well as all the Spectra Optics filters, has come from Mexican mines represented by Gary Anderson. Karl Lambrect Company sells both Mexican and South African calcite. We have obtained a sample of the South African material and it is of significantly higher whiteness of any previous Mexican calcite. Lambrect claims to have a significant stock of such material and guarantees its quality at $2000 1 kg.

Harry Ramsey has had extensive conversations with representatives or owners of calcite mines and sample calcite has been examined. If high quality calcite exists, there is no reason to look further for alternative materials. It would appear prudent therefore to have someone knowledgeable about calcite and calcite working actually visit and examine to known mined stock piles to assess the claims of the mine owners and their representatives. Most critical is the stock of mined South African calcite, since this material, because of political problems, could rapidly become unavailable. In view of the high quality of the samples in hand, it is my opinion that sufficient calcite can be obtained.

2. Metal Acid Pthalates

The most attractive alternate crystal appears to be potassium acid pthalate (KAP) or one of its alkali metal relatives. These crystals have been grown in fairly large sizes and have an effective birefringence nearly that of calcite. We have some of the crystals produced at New Mexico State

and found them to be of non-uniform birefringence. These crystals are water soluable, fragile, and extremely sensitive to thermal shock.

At present, I feel that investigating KAP and its relatives is a lower priority than the plastic materials discussed below.

3. Plastic Materials

Polyvinyl alcohol (PVA) is a major commercial plastic. It is the plastic film base material for Polaroid Corporation polarizer and waveplate sheets. PVA becomes birefringent when stretched. Reasonable stretches yield birefringencies up to three times quartz. The plastic can be stretched very uniformly. The sheet material can be laminated to form birefringent elements. We have made a 80Å tunable filter using PVA sheet. The sheet and the laminate are scatter free and are obtainable in at least 15 cm diameter.

PVA sheet should be considered for the purpose of making tunable prefilters for the final filter. Solc type filters become much more attractive using laminate sheets. I would recommend looking into the Universal blocking filter using plastic sheet.

4. Michelson Elements as Analog Lyot Elements

Evans in 1949 suggested that Michelson interferometers could be used to build Lyot filters. The response of a fixed Michelson is identical to that of a Lyot element. A fixed wide field Michelson offers two significant advantages for elements narrower than 1/4 Å. First, since a Michelson is roughly cubical regardless of its effective FWHM for elements narrower than 1/4 Å, there is increasing greater length advantages. Second, a wide field Michelson has a field of view which is the square of the field of an identical bandpass wide field Lyot element.

We have built a prototype 1/20 Å polarizing wide field Michelson. The device is quite stable and easily tuned by rotation of a waveplate or polarizer. a 1/50 Å filter using 2 or 3 wide field Michelsons is a reasonably straightforward development.

In my opinion; a 1/50 Å tunable filter could almost eliminate the need for high dispersion spectrographs. It would provide an enormously powerful tool for line profile analysis and magnetic field strength.

FIGURE CAPTIONS

Figure 1 Pallet random vibration levels.

Figure 2 Filter random vibration levels.

Figure 3 Fringe patterns visible in preliminary test unit due to separation of waveplates.

Figure 4 Mechanical drawing of Mark I and II Preliminary Test Unit (dimensions in cm).

Figure 5 Photograph of Mark II Preliminary Test Unit and mounting fixture for shake table.

Figure 6 Mechanical drawing of prototype filter module (dimensions in cm).

Figure 7 Photograph of prototype filter module prior to assembly.

Figure 8 Exploded photograph of prototype filter module and sealed enclosure.

Figure 9 Photograph of assembled sealed enclosure.

Figure 10 Transmission versus number of modules for several values of reflection loss per surface.

Figure 11 Reflection coefficient (full scale one percent) versus wavelength for a simple four layer A-R coating. Properties calculated using a published (Thetford, Optica Acta 11, 159, 1969) design and program thin film.

Figure 12 Drawing of Halle oven showing location of the quartz probe and controller sensors.

Figure 13 Optical schematic for the thermal controller test system.

Figure 14 Temperature response of the quartz probe (oven) and calcite interior versus time.

Figure 15 Temperature sensitivity and fringes/cm/^{o}C versus wavelength at 40^{o}C.

Figure 16 Free spectral range versus wavelength for a one centimeter calcite element.

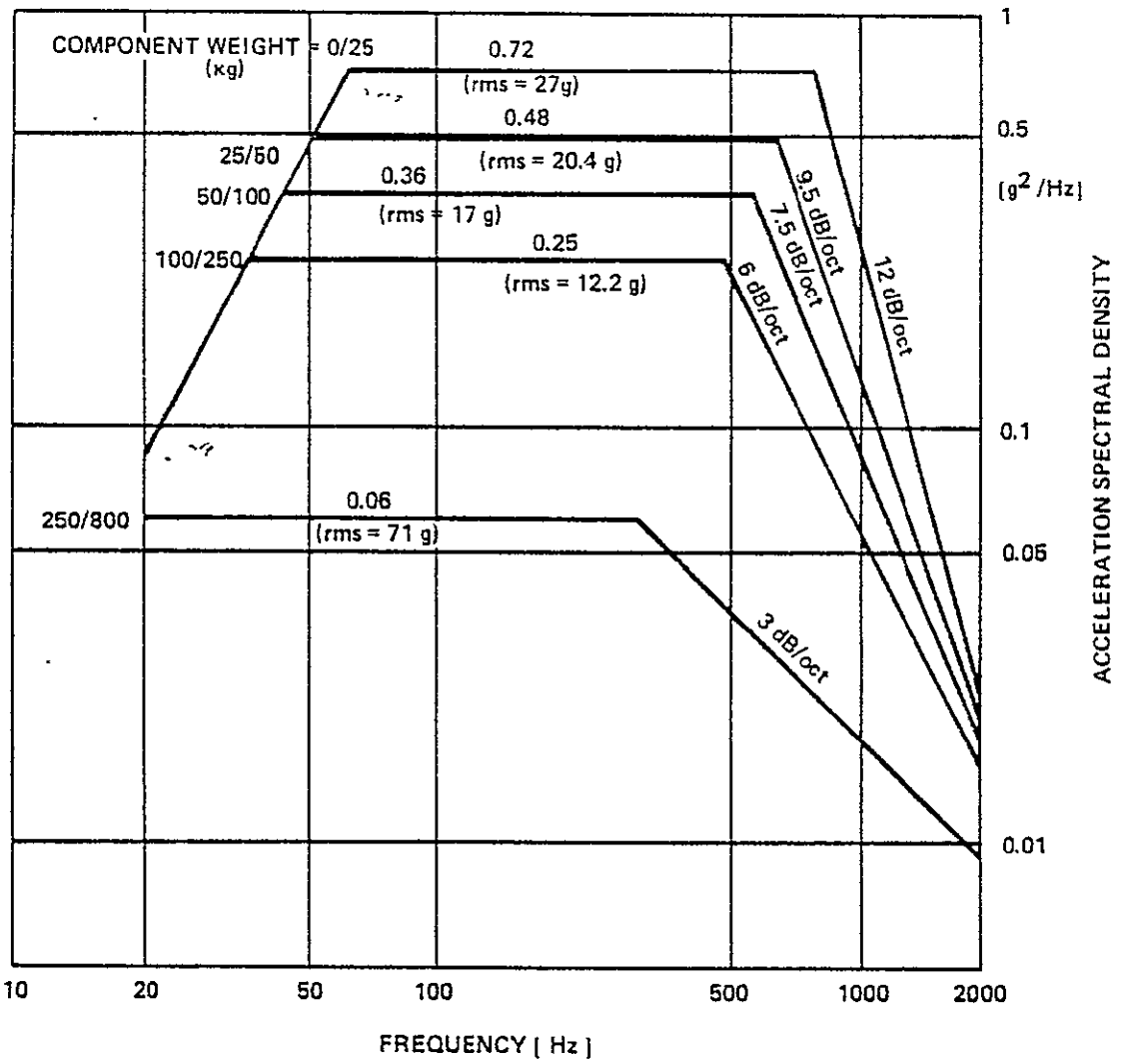
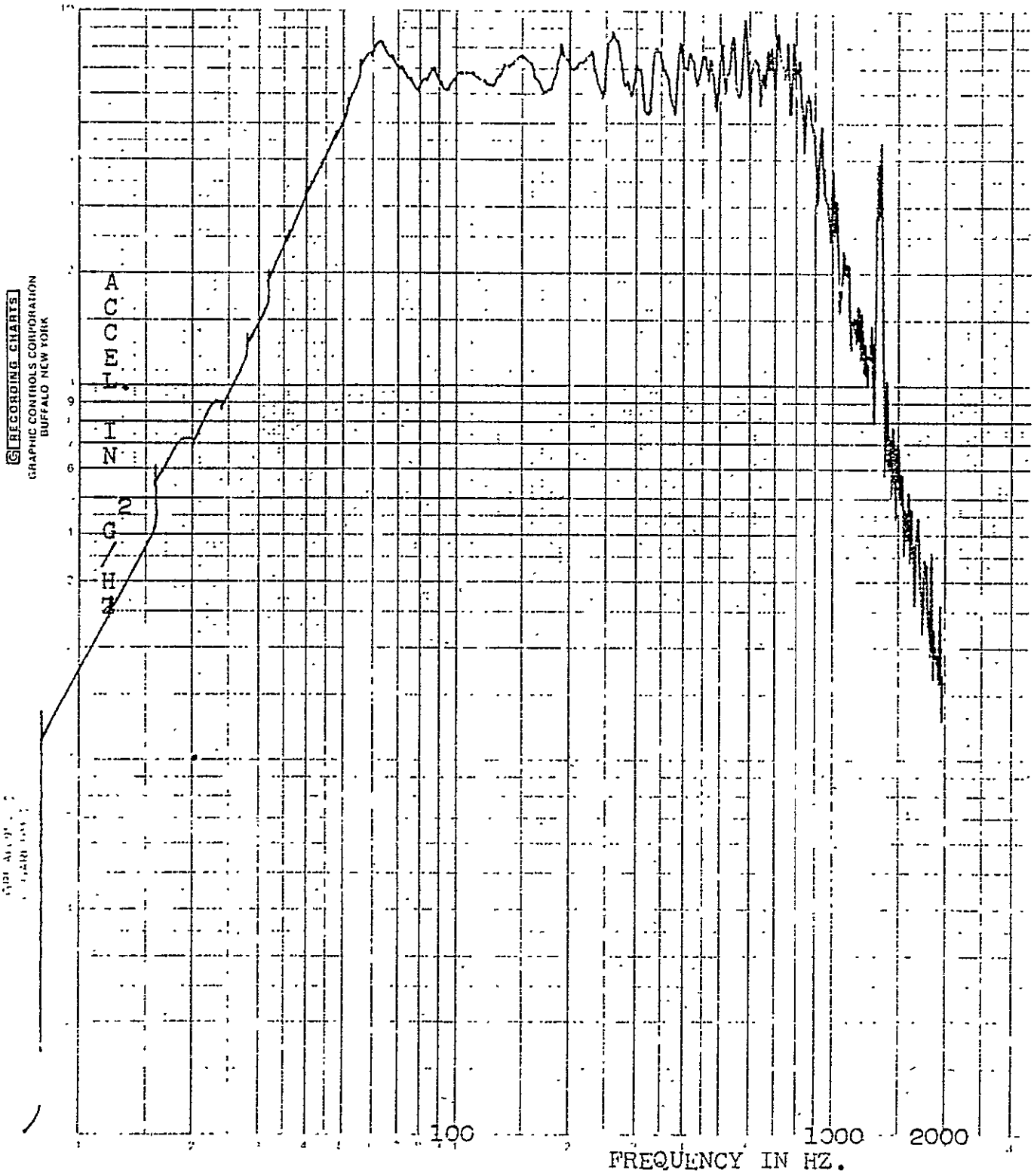


FIGURE 1 PALLET RANDOM VIBRATION LEVEL

RUN NO.- 6 DATE 3-17-76 TIME- 1107
FIXTURE [ ]
SPECIMEN [X] AXIS- Z GRMS- 27.2
I.D.- Refrig Filter

F.S.- 1.0 G^2/HZ



(31)

Figure 2

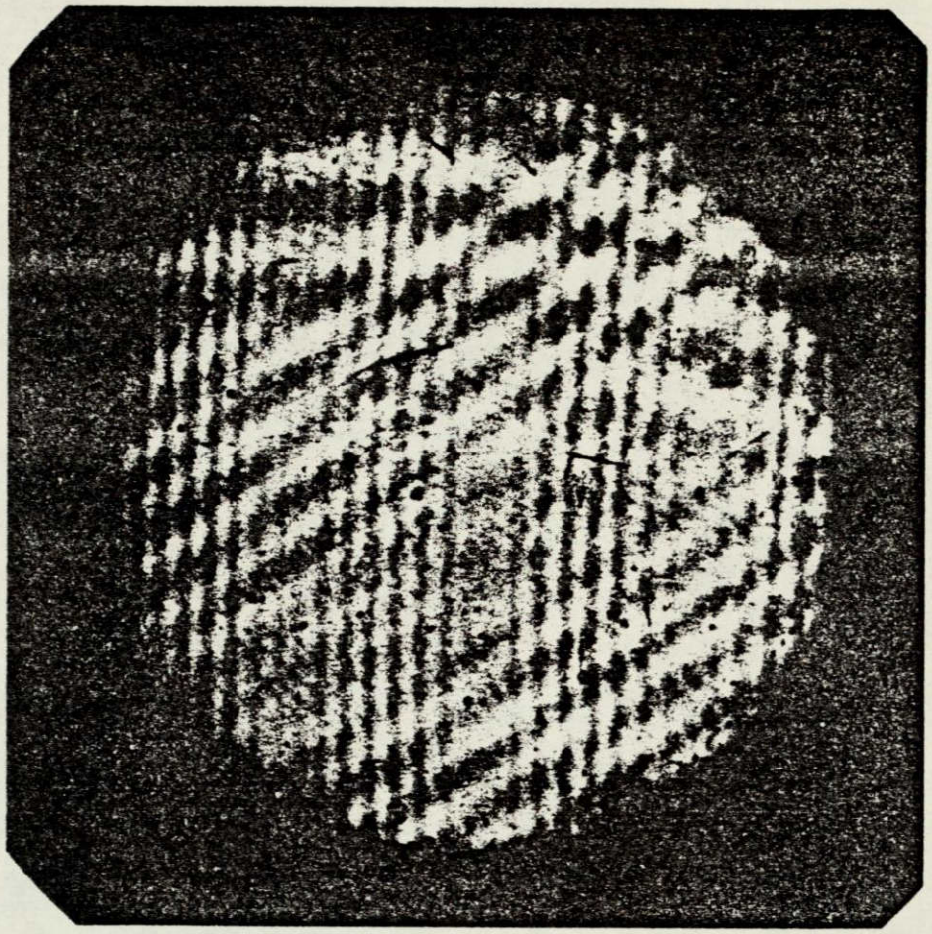


Figure 3
B-33

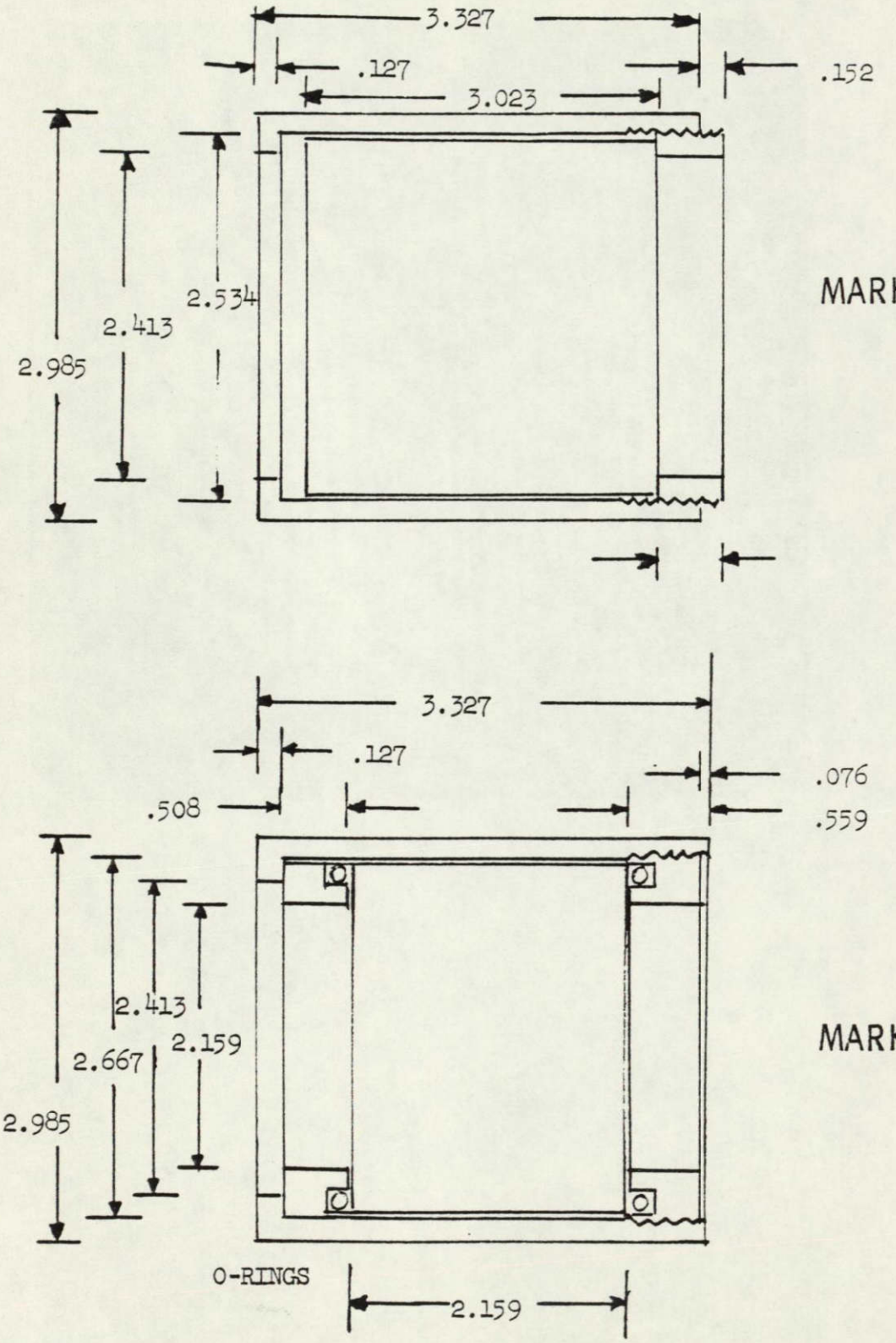
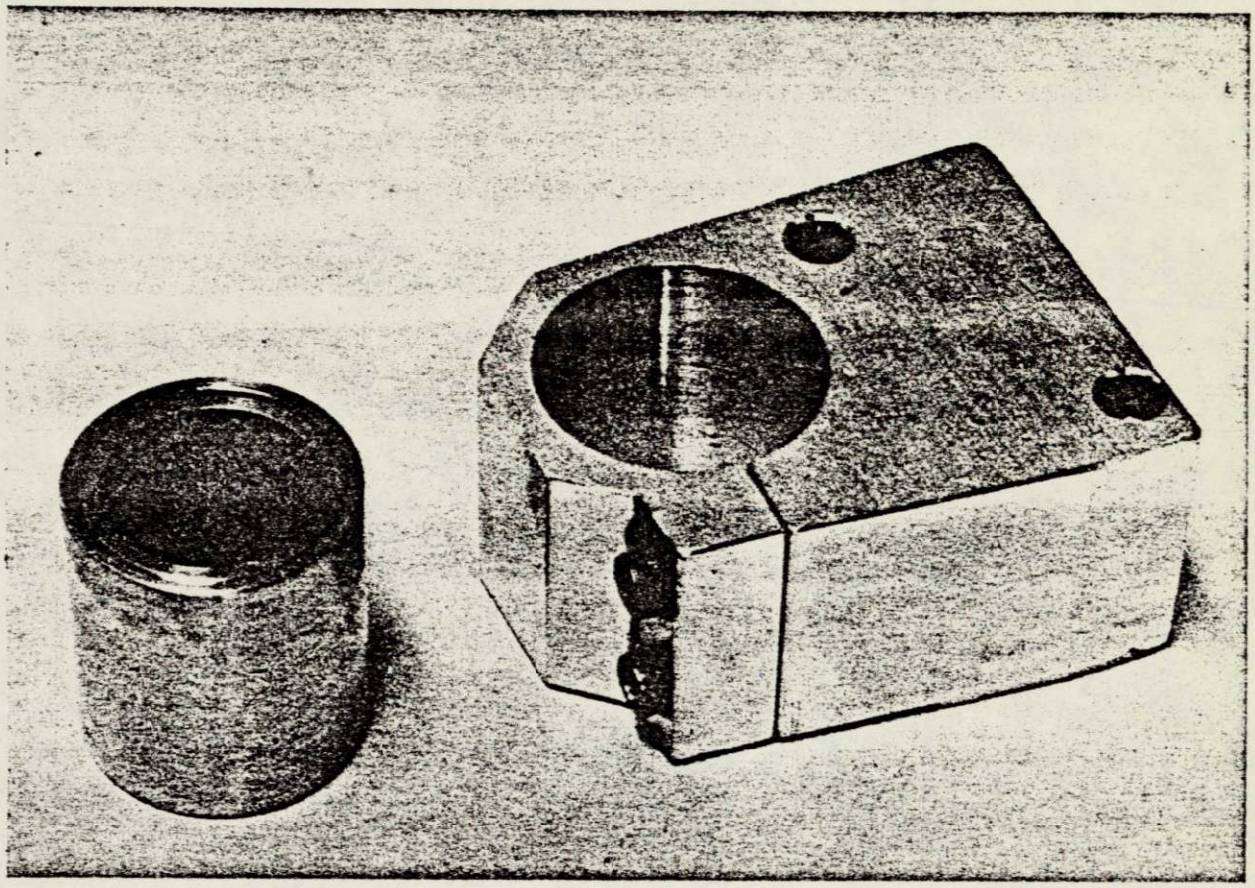

3.327
.127
3.023
.152
2.534
2.413
2.985
MARK I
3.327
.127
.508
.076
.559
2.413
2.667
2.159
2.985
MARK II
O-RINGS
2.159

Figure 4
B-34



C-3

Figure 5
B-35

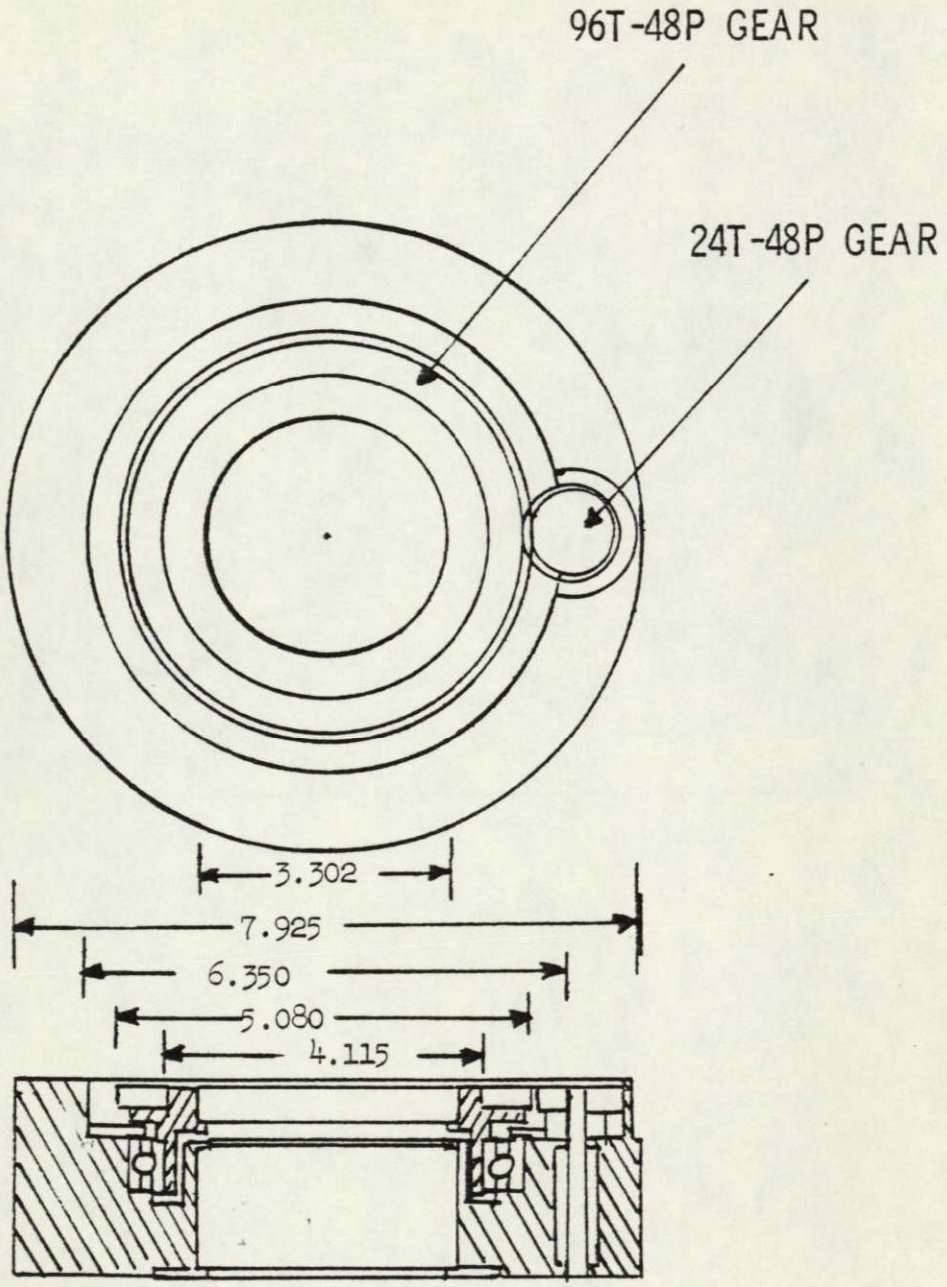


Figure 6

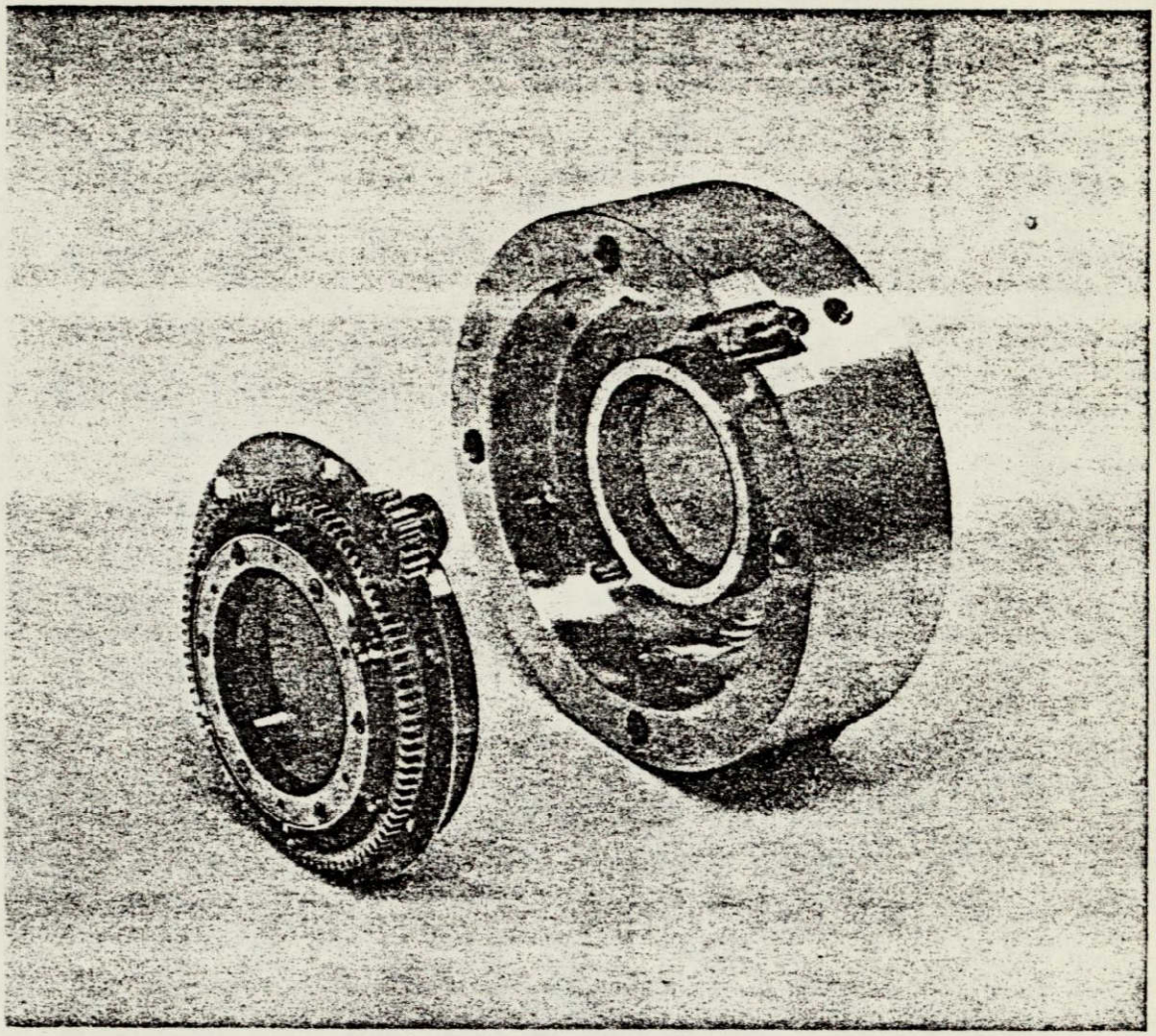


Figure 7

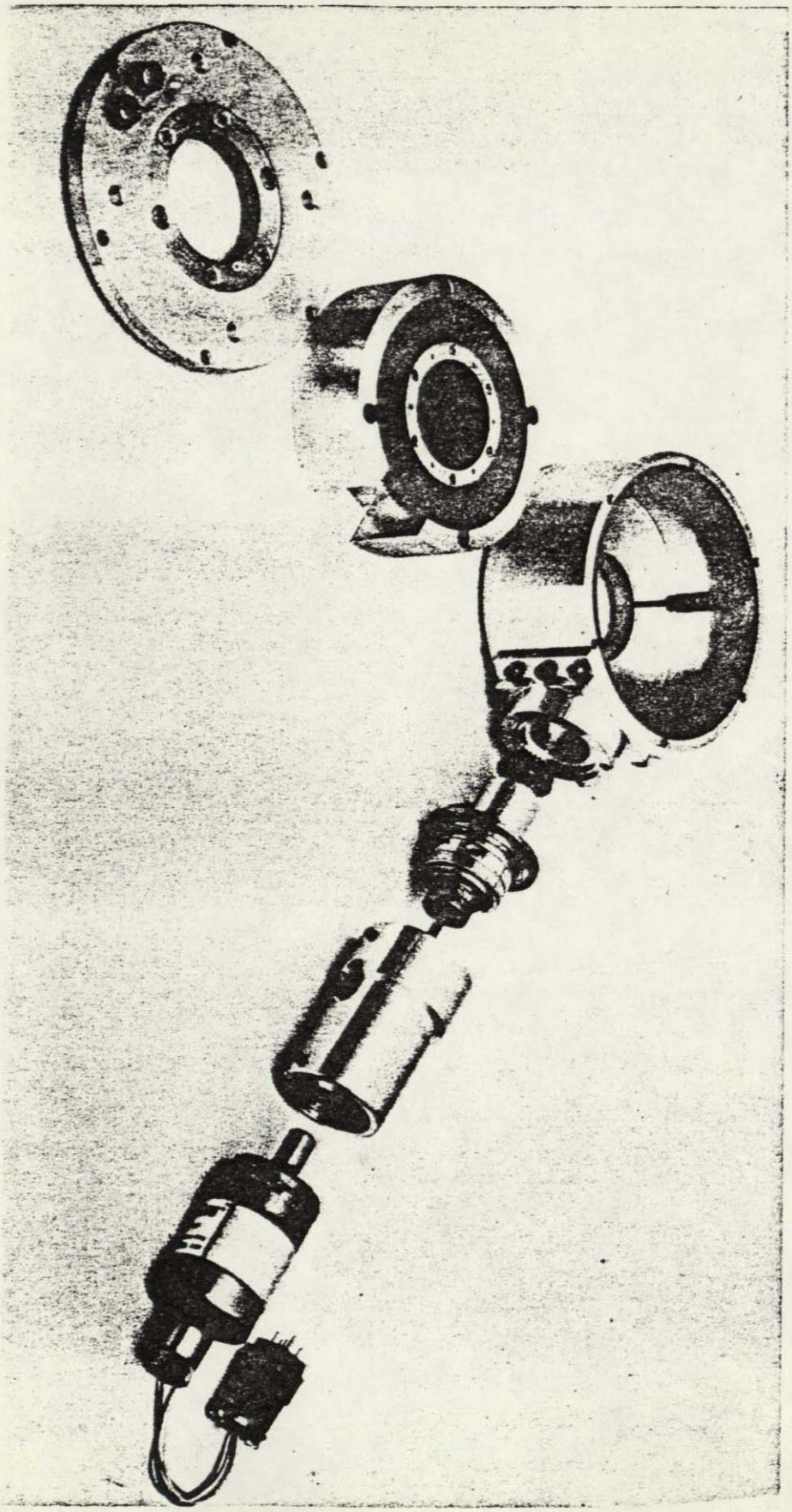


Figure 8

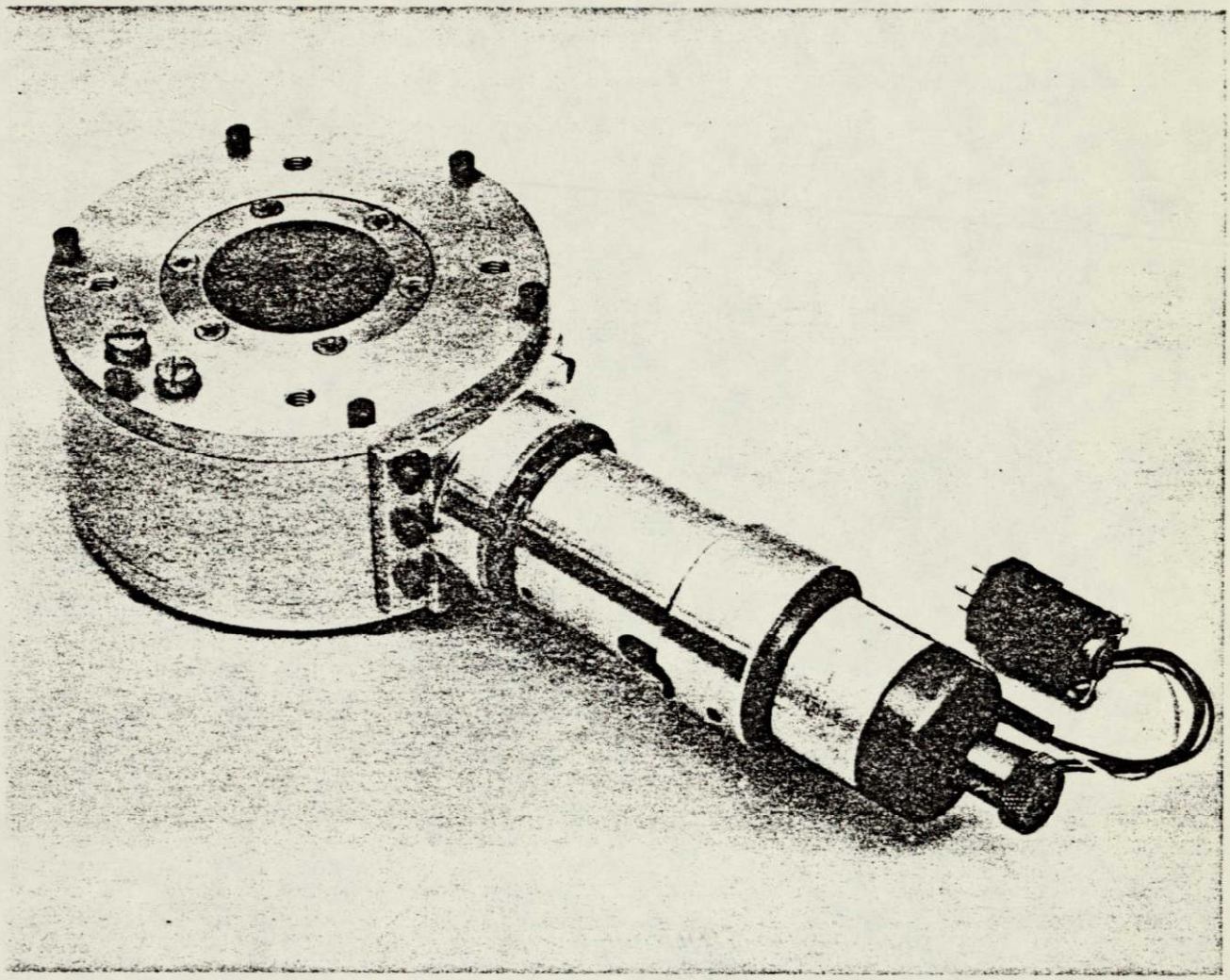


Figure 9

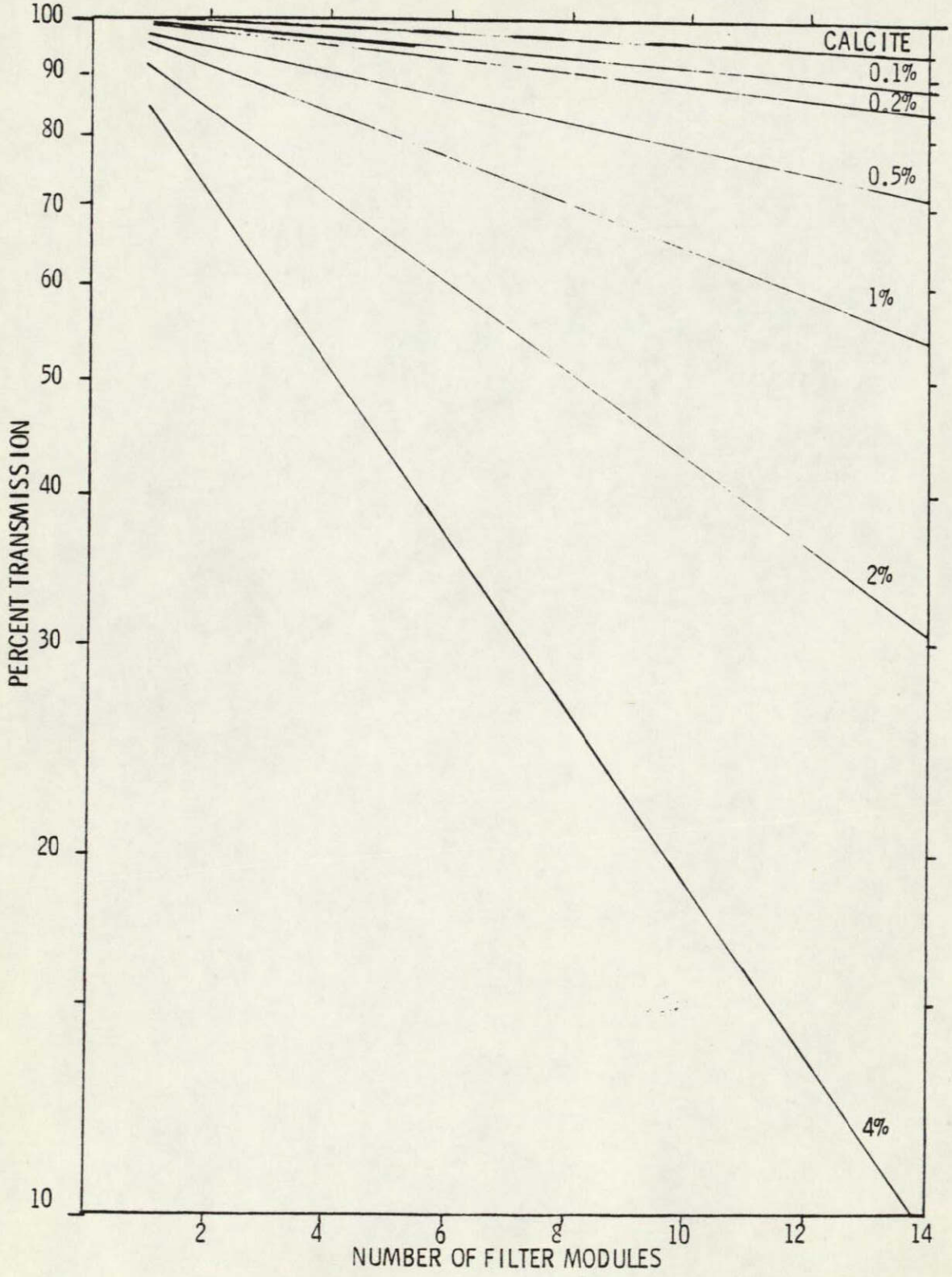

CALCITE
0.1%
0.2%
0.5%
1%
2%
4%
PERCENT TRANSMISSION
100
90
80
70
60
50
40
30
20
10
2
4
6
8
10
12
14
NUMBER OF FILTER MODULES

Figure 10

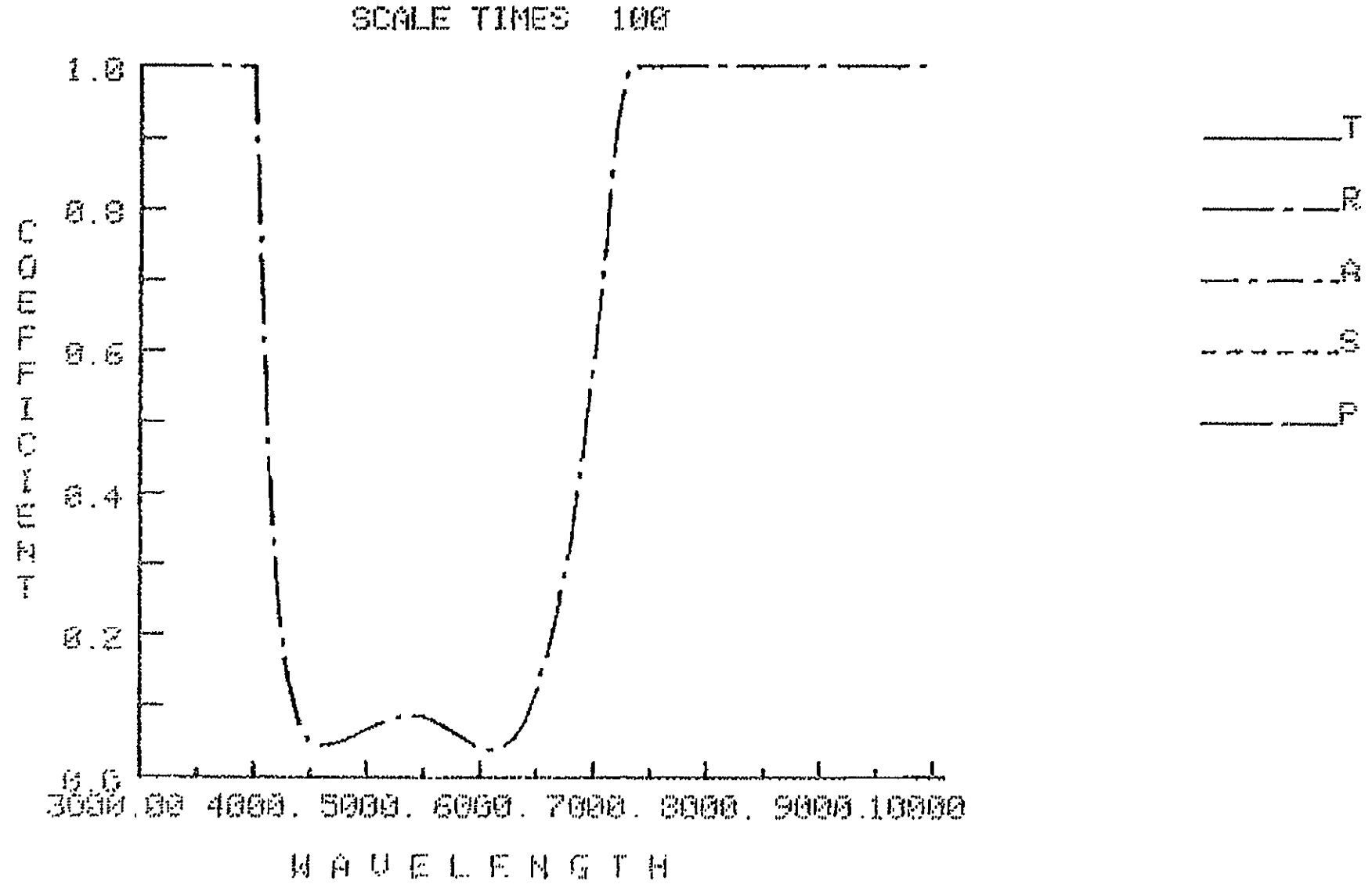


Figure 11
B-41

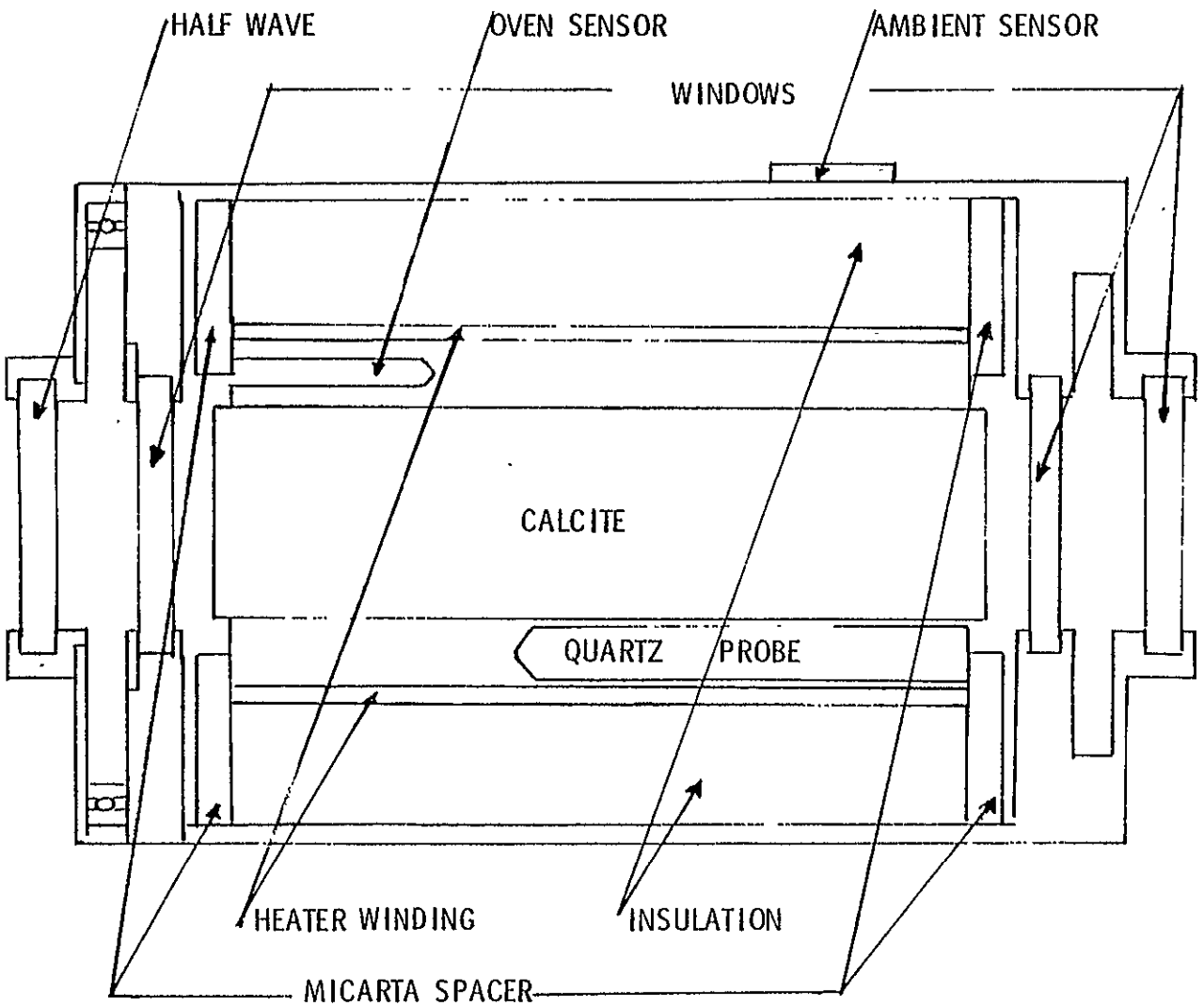
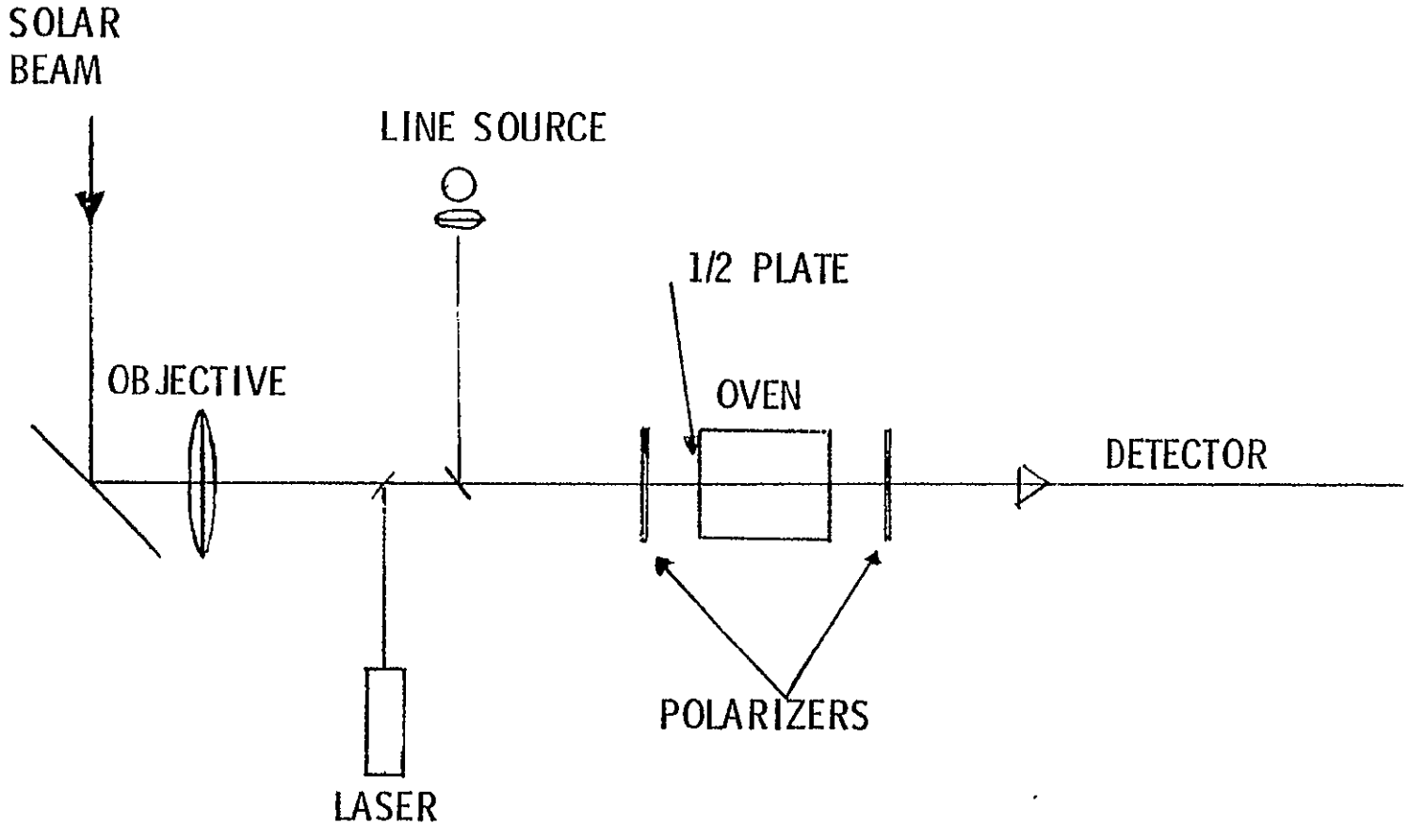


Figure 12
B-42



OPTICAL SCHEMATIC

Figure 13

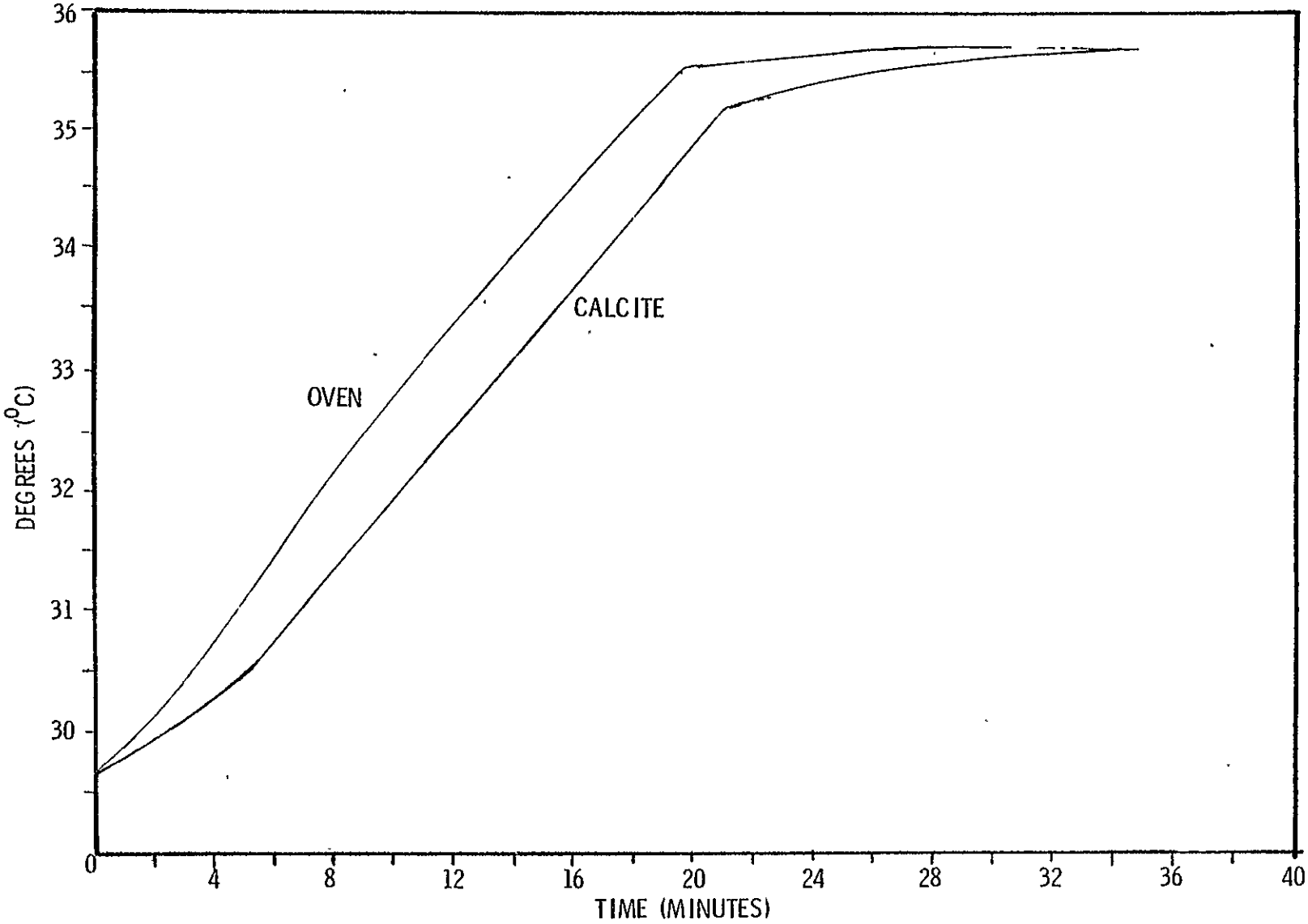

36
35
34
33
32
31
30
0
DEGREES (°C)
OVEN
CALCITE
4
8
12
16
20
24
28
32
36
40
TIME (MINUTES)

Figure 14

Figure 15
B-45

FREE SPECTRAL RANGE FOR ONE CM OF CALCITE
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
3000
4000
5000
6000
7000
8000
9000
10000
11000
12000
WAVELENGTH

Figure 16