OPTICS OF A BIREFRINGENT PLATE WITH APPLICATIONS TO ELLIPSOMETRY

by

Dale Arthur Holmes

Prepared under Contract No. NAS 8-5269 by

Electrical Engineering Department

Carnegie Institute of Technology

Pittsburgh, Pennsylvania 15213

for

National Aeronautics and Space Administration

OPTICS OF A BIREFRINGENT PLATE WITH APPLICATIONS TO ELLIPSOMETRY

by

Dale Arthur Holmes

Prepared under Contract No. NAS 8-5269 by
Electrical Engineering Department
Carnegie Institute of Technology
Pittsburgh, Pennsylvania 15213
for
National Aeronautics and Space Administration

TABLE OF CONTENTS

LIST OF FIGURES

GLOSSARY

a	Angle between analyzer transmission axis and compensator fast axis. Fig. F-1.
a	Subscript identifying a quantity calculated by geometrical optics.
a_o	$1/n_o^2$.
$[a]$	Index tensor.
$a_{11}, a_{12}, \ldots$	Elements of index tensor.
a_{xx}, a_{yy}, a_{zz}	Elements of diagonalized index tensor.
b, d, l, h	Physical dimensions. Fig. 4-6.
c	Azimuth of fast axis. Fig. F-1.
d, d_2, d_3	Plate thickness.
d, b, h, l	Physical dimensions. Fig. 4-6.
e	Angle between linearly polarized electric field direction and compensator fast axis. Fig. F-1.
e	Subscript identifying a quantity calculated by exact theory.
E	Modulus of elasticity.
E	Electric field.
$\bar{\mathcal{E}}_q^+, \bar{\mathcal{E}}_q^-$	Electric field vectors in qth medium.
$\bar{E}_q^+, \bar{E}_q^-$	Electric field amplitude vectors (not functions of spatial coordinates or time).
F	Magnitude of force. Fig. 4-6.
$\bar{\mathcal{H}}_q^+, \bar{\mathcal{H}}_q^-$	Magnetic field vectors in qth medium.
i	Angle of incidence.
i	Subscript identifying quantity with incident wave.

I	Intensity of light beam.
k, n	$n-jk$ is refractive index of plate in Fig. 3-1.
K, N	$N-jK$ is refractive index of substrate in Fig. 3-1.
k_c, n_c	Calculated values of k and n .
l, m, n	Direction cosines.
l_v, m_v, n_v	Direction cosines.
$n_{\alpha q}, n_{\beta q}, n_{\gamma q}$	Principal refractive indices in qth medium. $n_{\beta q} = n_{yq}$.
n_1, n_2, n_3	Refractive indices in media 1, 2, and 3. Fig. 3-1.
n_c, k_c	Calculated values of n and k.
n_i	Refractive index of surrounding isotropic medium.
n_q^+, n_q^-	Refractive indices of traveling waves in qth medium.
n_o	Refractive index of unstressed crystal. $n_o = n_2 = n - jk$.
n_{xx}, n_{yy}, n_{zz}	Principal refractive indices in stressed crystal.
N, K	$n_3 = N - jK$. Substrate index. Fig. 3-1.
o	Subscript identifying quantities associated with output wave.
p	Subscript identifying quantities associated with electric field parallel to the plane of incidence.
p	Angle defining location of polarizer transmission axis.
q	Subscript denoting medium under consideration.
q_{11}, q_{12}, q_{44}	Piezo-optic constants.
$\bar{r}$	Position vector.
r	Subscript identifying quantities associated with reflected wave.
$r_{12}, r_{23}, r_{34}, \ldots$	Two media amplitude reflectances.
R_p, R_s	Amplitude reflectances.

s	Subscript denoting quantities associated with electric field perpendicular to the plane of incidence.
$\hat{s}_q^+ , \hat{s}_q^-$	Unit wave normal vectors.
$t_{12}, t_{23}, t_{34}, \ldots$	Two media amplitude transmittances.
V	Magnitude of uniaxial stress.
x, y, z	Coordinate axes.
α	Absorption coefficient. $\alpha = 4\pi k / \lambda$.
$\alpha_q, \beta_q, \gamma_q$	Principal dielectric axes in qth medium.
δ_q^+ , δ_q^-	$\delta_q^+ = 2\pi n_q^+ \cos\theta_q^+ / \lambda$, $\delta_q^- = 2\pi n_q^- \cos\theta_q^- / \lambda$.
λ, λ_o	Vacuum wavelength.
ϕ_q	Angle between α_q axis and x axis. Fig. I-1.
ρ	Complex number defining polarization state of wave. Appendix F.
θ_q^+ , θ_q^-	Angle between $\hat{s}_q^+$ ($\hat{s}_q^-$) and positive z axis. Fig. I-1.

ACKNOWLEDGEMENTS

Throughout the course of this research, I have benefited from many valuable discussions with Prof. D. L. Feucht and Prof. R. L. Longini; their support and encouragement have made this work possible. I am indebted to Prof. Longini for suggesting the research problem of investigating the effect of an externally applied mechanical strain on the optical properties of single crystal GaAs through the use of an ellipsometer. Mr. E. Litot gave invaluable assistance in the mechanical design and construction of the ellipsometer. Mr. J. Valerio assisted in various tasks pertaining to electronic detection circuitry and optical calibration of the ellipsometer. I am grateful to the National Aeronautics and Space Administration for financially supporting this work through Carnegie Institute of Technology Contract NAS8-5269.

ABSTRACT

The ellipsometric measurement of the optical properties of mechanically stressed single crystals of GaAs is the central topic under consideration. GaAs becomes anisotropic when stressed, therefore, the crystal optics problem of electromagnetic wave reflection from a birefringent specimen is theoretically investigated using the Maxwell equations.

The theoretical description of optical compensators is critically examined. It is found that the geometrical optics description is inadequate when retardation plates and rotary compensators are used for precise measurements in well collimated, monochromatic light. A very general method for using a Senarmont compensator in ellipsometry is outlined; the method is applicable, without modification, when the compensator is inexact, when multiple internal reflections are present, and even when the compensator is dichroic. A new type of retardation plate, which uses an isotropic rather than an anisotropic plate, is proposed.

Detailed consideration is given to the methods for determining the optical constants of unstressed GaAs and the piezo-optical constants of stressed GaAs through the use of ellipsometry data, including a discussion of how experimental errors are propagated in the derived inversion relations.

An infrared ellipsometer with a specimen chamber designed so that stress (or strain) can be applied to a GaAs specimen and using a GaAs injection laser source is described.

INTRODUCTION

The original objective of this work was to study the effects of strain on the performance of GaAs p-n junction lasers. The strain destroys the cubic symmetry of the GaAs crystal and the band structure is thus perturbed from the unstressed state. With the application of strain, one might expect a shift in the emission spectrum due to a change in the bandgap. The strain could force the transition probabilities to become dependent upon the propagation direction and polarizations, which in turn would effect the threshold and power output characteristics of the laser. After this research was begun, two reports appeared describing the effect of strain on GaAs lasers[a,b]. These brief reports of experimental work showed that it was possible to frequency tune the laser, lower its threshold current, and increase its output for a given current input through the use of an appropriately applied strain; no attempt was made to correlate the experimental results with quantitative theoretical predictions.

The probability per unit time for a transition between a given state in the valence band and a given state in the conduction band is the same for both stimulated emission and absorption of a photon, therefore it is possible to derive criteria for laser action if optical absorption data is available for material corresponding to the type to be used in a proposed laser[c]. In order to make quantitative predictions concerning the behavior of a GaAs laser under strain, one possible approach is then to first investigate the optical properties of laser-like GaAs single crystals under strain, with particular interest in the strain-induced anisotropy in the absorption characteristics. This approach was adopted and this work describes the progress that has been made to date.

---

[a] F. Ryan, R. Miller, Appl. Phys. Lett. 3, 162 (1963).

[b] D. Meyerhoffer, R. Braunstein, Appl. Phys. Lett. 3, 171 (1963).

[c] W. P. Dumke, Phys. Rev. 127, 1559 (1962).

In Chapt. I the interaction of an electromagnetic wave with an anisotropic crystalline plate is discussed. The geometry postulated is considered to be sufficiently general to describe the transmission and reflection characteristics of a GaAs specimen for a variety of crystallographic orientations and strain states. It turns out that the solution to the specific problem considered in Chapt. I represents a new contribution in itself. The results of Chapt. I are used to refine our understanding of the selected measurement technique in Chapt. 2. I might remark here that an ellipsometric measuring method was adopted. In an ellipsometric measurement two quantities, phase difference and amplitude ratio, are measured, hence, an ellipsometric measurement, in principle contains more information than a conventional transmission type measurement wherein only one quantity, the transmitted intensity, is measured. In addition, the ellipsometric measurement does not depend on the intensity of the light being measured, therefore, no severe restrictions are placed on the stability (with respect to intensity) of the light source. Chapt. 3 discusses how the optical properties of the GaAs specimen can be calculated after measuring the polarization state of the incident and reflected light. Chapt. 4 is an description of the present state of the ellipsometer.

Most of the chapters present mainly the important ideas with numerical examples and detailed derivations generally left to the appendices. Apps. B,C, and D, appropriate to Chapt. 2, are relatively self-contained works and can be read without reference to Chapt. 2. The notation in these appendices differs, in places, from that used in the text. A bibliography of unreferenced works related to various topics discussed in this work can be found immediately following the references.

CHAPTER I

Electromagnetic Wave Interaction with a Birefringent Plate

Problems involving reflection and refraction at planar boundaries of anisotropic media are, in general, very complicated; in fact, the expressions describing reflection from and transmission through a birefringent plate have never, to the best of my knowledge, been published for the completely general case of arbitrary orientation of the principal dielectric axes of the plate with an oblique angle of incidence*. The approach to this problem has been somewhat evolutionary with each new publication treating a more general situation than those previously. In this chapter I carry the evolutionary process a bit further by considering a problem more general than any found in the literature.

I.1 DESCRIPTION OF THE PROBLEM

The birefringent plate geometry is shown in Fig. I-1. A right hand xyz coordinate system is located such that the xy plane coincides with the interface between media 1 and 2. In each of the three anisotropic media (The media surrounding the plate are also considered to be birefringent.) one of the principal dielectric axes, β_q , is parallel to the y axis, while the orientation of the other two axis, denoted by the α_q and γ_q , $q = 1,2,3$, directions, is arbitrary. Because of the orthogonality of the principal dielectric axes, all of the $\alpha_q \gamma_q$ planes are parallel to the xz plane. The angle between the α_q direction and the x direction is ϕ_q . Each medium is biaxial, with principal refractive indices $n_{\alpha q}$, $n_{\beta q} \equiv n_{yq}$ and $n_{\gamma q}$. The birefringent plate has thickness d .

---

* Throughout this work, "anisotropic" and "birefringent" will be used interchangeably to characterize a medium which is nonmagnetic, nonoptically active, homogeneous, and lossless, except when specified otherwise, and for which the relative dielectric constant or refractive index is a tensor quantity.

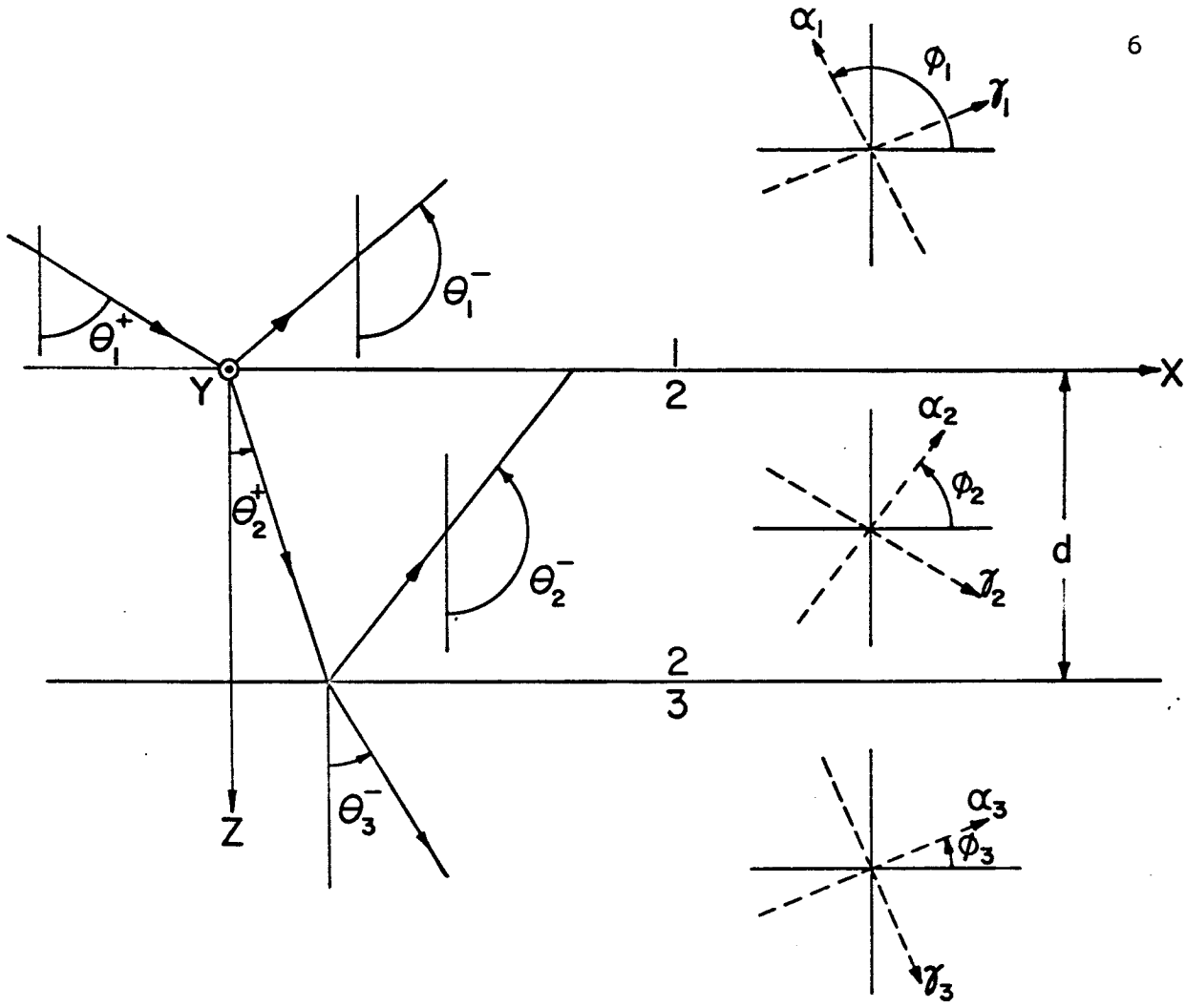


Fig. I-1 Diagram of a birefringent plate surrounded on both sides by semi-infinite birefringent media. The plane of incidence is the plane of the paper. For each medium, one of the principal dielectric axes is perpendicular to the plane of incidence. The subscripted α and γ symbols denote the directions of the other two principal axes. Along a principal axis α_q the principal refractive index is $n_{\alpha q}$ while along γ_q the principal index is $n_{\gamma q}$, $q = 1,2,3$. Since the β_q axes (not shown) are parallel to the y axis in each medium, the refractive index along a β_q axis shall be denoted by n_{yq}. The angle ϕ_q is the angle between the α_q direction and the x direction. The various subscripted θ quantities are angles between unit wave normal vectors and the positive z axis. The birefringent plate (medium 2) has thickness d. All media are assumed homogeneous, nonoptically active and nonmagnetic. The plate surfaces are assumed flat and parallel.

An electromagnetic wave of the form*

$$\bar{\mathcal{E}}_1^+ = \bar{E}_1^+ \exp\left\{ j\omega t - j(2\pi n_1^+/\lambda)\, \hat{s}_1^+ \cdot \bar{r} \right\} \qquad (1)$$

is incident from medium 1. In (1), ω is the angular frequency, t is time, λ is the vacuum wavelength, $\bar{r}$ is a position vector, n_1^+ is the refractive index of the incident wave and $\hat{s}_1^+$ is the unit wave normal vector. The unit wave normal $\hat{s}_1^+$ has an x component $\sin\theta_1^+$ and a z component $\cos\theta_1^+$, where θ_1^+, commonly called the angle of incidence, is the angle between $\hat{s}_1^+$ and the positive z axis. Note that the plane of incidence is parallel to the xz plane. For a given value of θ_1^+, two waves can propagate in the $\hat{s}_1^+$ direction[1],** each having a different value of n_1^+.

Because of the manner in which the plane of incidence and the directions of the principal dielectric aces have been specified, one of these waves will propagate in an ordinary fashion, that is, its refractive index is independent of θ_1^+ and, in fact, is given by $n_1^+ = n_{y1}$. The electric field vector of this "ordinary" wave will be perpendicular to the plane of incidence. The incident ordinary wave will experience ordinary reflection and refraction at both the $z = 0$ and the $z = d$ interfaces, that is, all of the waves excited in media 1,2, and 3 by the incident ordinary wave will have refractive indices which are independent of their respective directions of propagation. The ordinary waves in the qth medium will have refractive index n_{yq}. The net result here is that the equations describing reflection from and transmission through the plate will be identical to those derived for isotropic media.

The other type of wave, called the extraordinary fashion, which can travel in the θ_1^+ direction, propagates in an extraordinary fashion, that is, its refractive index n_1^+ is a function of θ_1^+ and is given by*

$$\left(n_1^+/n_{x1}n_{z1}\right)^2 = \left[n_{x1}^2 \sin^2(\theta_1^+ - \phi_1) + n_{z1}^2 \cos^2(\theta_1^+ - \phi_1) \right]^{-1}. \qquad (2)$$

---

* A symbol with a bar, such as $\bar{E}_1^+$, denotes a vector quantity while a broken bar indicates a unit vector, such as $\hat{s}_1^+$.

** Ref. 1 reviews many topics in the optics of anisotropic media.

*** The derivation of (2) is considered later.

The electric vector of the extraordinary wave has only x and z components, i.e., is contained wholly in the plane of incidence; furthermore, the electric vector, in general, is not perpendicular to $\hat{s}_1^+$. The extraordinary incident wave will experience extraordinary refraction and reflection at both the $z = 0$ and $z = d$ interfaces. In this chapter, I will give the formula for the (complex) amplitude of the wave reflected in medium 1 and of the wave transmitted into medium 3, assuming an incident wave of the extraordinary kind.

Now that the geometry of Fig. 1.1 has been described, it is convenient to discuss some of the already published works. The case where media 1 and 3 are isotropic and $\phi_2 = 0$ has been very comprehensively treated by Schopper[2]. The reflection properties of a geometry for which medium 1 is isotropic and $\phi_2 = \phi_3 = 0$ are summarized by Winterbottom[3]. Two workers[4] have investigated a geometry which is, at the same time, both more and less general than that of Fig. 1.1. It is more general in the sense that the orientation of the principal dielectric axes of the plate is completely arbitrary and less general in the sense that media 1 and 3 are considered isotropic and only normal incidence is considered.

1.2 SOLUTIONS FOR EXTRAORDINARY WAVES

In medium 1 there is a reflected wave traveling in a direction $\hat{s}_1^-$ with refractive index n_1^-, whose electric field can be represented by

$$\bar{\mathcal{E}}_1^- = \bar{E}_1^- \exp\{j\omega t - j(2\pi n_1^-/\lambda)\, \hat{s}_1^- \cdot \bar{r}\}\ , \tag{3}$$

while in medium 3 there is a transmitted wave traveling in a direction $\hat{s}_3^+$ with refractive index n_3^+, whose electric field can be represented by

$$\bar{\mathcal{E}}_3^+ = \bar{E}_3^+ \exp\{j\omega t - j(2\pi n_3^+/\lambda)\, \hat{s}_3^+ \cdot \bar{r}\}\ . \tag{4}$$

Within the plate (medium 2), multiple internal reflections are set up which can be characterized by a single "positively" traveling wave with wave normal $\hat{s}_2^+$ and refractive index n_2^+ and a single "negatively"

traveling wave* with wave normal $\hat{S}_2^-$ and refractive index n_2^-. If θ_q^+ (and θ_q^-) are defined as angles between $\hat{S}_q^+$ (and $\hat{S}_q^-$) and the positive z axis, then both $\hat{S}_q^+$ and $\hat{S}_q^-$ have x components $\sin\theta_q$ and z components $\cos\theta_q$, where the + and - superscripts are understood. The various θ_q^+ and θ_q^- are shown in Fig. I-1. The waves in medium 2 have electric fields which can be written analogously to (1), (3), and (4).

As mentioned earlier, the extraordinary waves under consideration have only x and z components of electric field. The reflection and transmission properties of the plate shall be characterized in terms of the tangential electric fields. Thus, if E_{x3}^+ (E_{x1}^-) is the amplitude of the x component of transmitted (reflected) electric field we have, from App. A,

$$\frac{E_{x3}^+}{E_{x1}^+} = \frac{t_{12}\,t_{23}\exp(j\delta_3^+ d)}{\exp(j\delta_2^+ d) - r_{21}\,r_{23}\exp(j\delta_2^- d)}\,, \tag{5}$$

$$\frac{E_{x1}^-}{E_{x1}^+} = \frac{r_{12}\exp(j\delta_2^+ d) + r_{23}(t_{12}+r_{21})\exp(j\delta_2^- d)}{\exp(j\delta_2^+ d) - r_{21}\,r_{23}\exp(j\delta_2^- d)}\,. \tag{6}$$

The subscripted r and t quantities are two media tangential reflectances and transmittances. For example, if media 2 and 3 are identical, then r_{21} would have the following meaning: If a source wave, traveling in a direction $\hat{S}_2^-$, is incident (from medium 2) upon the interface $z = 0$, then the reflected wave in medium 2, traveling in a direction $\hat{S}_2^+$, has an x component of electric field given by $E_{x2}^+ = r_{21} E_{x2}^-$. The δ quantities are given by

$$\delta_2^+ = 2\pi n_2^+ \cos\theta_2^+ / \lambda\,,$$

with similar relations for δ_2^- and δ_3^+.

---

* My classifications, "positively" and "negatively" traveling, stem from the fact that, when $\theta_1^+ = 0$ (normal incidence), $\hat{S}_q^+$, $q = 1,2,3$ points in the positive z direction while $\hat{S}_q^-$ points in the negative z direction.

A more detailed discussion of the precise relations for the r and t quantities is contained in App. A. Also contained in App. A are the relations for the indices of refraction and for the unit wave normals of all the waves traveling in each medium. Eqtns. (5) and (6) have been written so that they resemble, to some extent, the corresponding solutions for isotropic media[5].

1.3 DISCUSSION

While I have not considered incident waves of the ordinary kind, I might remark that (5) and (6) can be adapted to this case if the various r, t, and δ quantities are appropriately defined. If the plate is absorbing and if the principal axes of the dielectric tensor coincide with the principal axes of the (effective) conductivity tensor[6], then the losses can be accounted for by assuming complex principal refractive indices, e.g., $n_{\alpha 2} \rightarrow n_{\alpha 2} - jK_{\alpha 2}$, where $K_{\alpha 2} > 0$. If the plate exhibits negative absorption[4,7], then the principal refractive indices would be complex quantities of the form $n_{\alpha 2} + jK_{\alpha 2}$ where $K_{\alpha 2} > 0$.

Since the various orientation angles ϕ_q are arbitrary, as are all of the principal indices of refraction, the solutions given in this chapter and in App. A can be adapted to a variety of problems involving electromagnetic wave interactions with anisotropic or isotropic plane-parallel layers. The solutions obviously can be used to describe the polarization state of light reflected from thin films or of light transmitted through crystalline plates. The latter subject, with particular emphasis on optical compensators, is treated in the next chapter. Some subtler applications might include the following. An analysis could be made of the multiple reflection effects in the air or cement filled gap in Glan-Foucault, Glan-Thompson, and other existing or proposed types of crystal prism polarizers[8]. Recent work on laser modulators[9], laser output couplers[10], adjustable microwave bidirectional couplers[11], and laser oscillators and amplifyers[7,12] could be extended to include the case of birefringent media. For those optical devices which are immersed in an

isotropic medium (air, for example) and which have $\phi_2 = 0$, one could predict the effects of misalignment of the principal axes by considering small values of ϕ_2.

The case of a birefringent plate surrounded by an isotropic medium is rather interesting. In this instance $|E_{x3}^+ / E_{x1}^+|^2$ represents the power transmittance. If we imagine, for the moment, that the plate is also an <u>isotropic</u> dielectric, then $\delta_2^+ = -\delta_2^-$ and, as shown by Collin[13], the incident wave will be wholly transmitted when $\delta_2^+ d$ is an integral multiple of π. In other words, the power transmittance is a periodic function of the plate thickness. Getting back to the anisotropic plate, it is apparent from App. A that $\delta_2^+ \neq -\delta_2^-$, in general, therefore, the periodicity of the power transmittance with plate thickness is radically different from the isotropic case.

In examining the transmission characteristics of birefringent plates, opticists have long been notorious for their use of the geometrical optics approximation. In this approximation, the multiple reflections are ignored. In (5), the effect of the internal reflections is eliminated by setting all of the two media reflectances equal to zero and all of the two media transmittances equal to unity, yielding

$$E_{x3}^+ / E_{x1}^+ = \exp\left(j[\delta_3^+ - \delta_2^+]\, d\right).$$

In short, the incident wave is wholly transmitted and the only effect of the plate is to impart a phase shift or, if the plate is lossy, the amplitude attenuation is a simple exponential function of plate thickness. A good example of the geometrical optics approach is contained in the work of Mandarino[14]. The geometrical optics approximation can be justified if the incident light contains a broad band of wavelengths or if the plate surfaces are not flat within a fraction of a wavelength and parallel. The impact of the laser technology has made these justifications questionable, however. Optical maser sources, such as the He-Ne laser, produce an extremely monochromatic and well collimated beam. A great number of research laboratories now have facilities for preparing paral-

lel-plate specimens with surface flatness tolerances of $\lambda/100$ and surface parallelism tolerances of 2 seconds of arc[15]. Even more critical is the fact that multiple internal reflections play a very important role in the buildup of electromagnetic waves in optical maser cavities, which are on the order of centimeters in length. Needless to say, a good part of the work to follow will deal with multiple internal reflections and some of their implications in optical devices and measuring techniques.

CHAPTER 2

Electromagnetic Aspects of Optical Compensators

Most optical compensators consist of one or more birefringent plates surrounded by an isotropic medium, usually air. If the two plate geometry of App. A is simplified to the extent that media 1 and 4 are identical and isotropic and media 2 and 3 have $\phi_2 = \phi_3 = 0$, (Fig. 1-1) we will still find that it can be adapted to describe the operation of most optical compensators. The transmitted wave is of interest when discussing compensators. This chapter will contain some preliminary discussion of the above mentioned simplified geometry as well as a discussion of the implications of multiple internal reflections in several types of optical compensators. A good part of the detailed discussion, including numerical examples, is contained in Apps. B, C, and D.

2.1 SIMPLIFIED GEOMETRY

When media 1 and 4 are both isotropic and identical, the distinction between ordinary and extraordinary waves vanishes; the incident, reflected and transmitted waves all have the same index of refraction regardless of the polarization of electric fields and, further, the unit wave normal for the transmitted wave is the same as that for the incident wave. It is still convenient to decompose the electric field vector of the incident and transmitted waves into two components, one wholly contained in the plane of incidence and the other perpendicular to the plane of incidence. The components in the plane of incidence are called p components while the components perpendicular to the plane of incidence are called s components. At oblique incidence, the p component is the vector sum of the x and z components and at normal incidence the p component is simply equal to the x component. For the incident and transmitted waves the p and s components are both perpendicular to the unit wave normal. Because the transmitted wave travels in the same direction as the incident

wave, we have that

$$E^+_{p4}/E^+_{p1} = E^+_{x4}/E^+_{x1} \;;\; E^+_{s4}/E^+_{s1} = E^+_{y4}/E^+_{y1} \;. \tag{7}$$

The p and s subscripts will be attached to various quantitues in this chapter to establish the component of incident (and transmitted) electric field under consideration.

For the simplified geometry, the results of Chapt. I and App. A can be reduced to

$$\begin{aligned}
&n^+_{p1} = n^+_{s1} = n^+_{p4} = n^+_{s4} \equiv n_i \;, \\
&\theta^+_{p1} = \theta^+_{s1} = \pi - \theta^-_{p1} = \pi - \theta^-_{s1} = \theta^+_{p4} = \theta^+_{s4} \equiv i \;, \\
&k^+_{p1} = -k^-_{p1} = k^+_{p4} \equiv k_{pi} = n_i/\cos i \;, \\
&k^+_{s1} = -k^-_{s1} = k^+_{s4} \equiv k_{si} = n_i \cos i \;, \\
&k^+_{pq} = -k^-_{pq} \equiv k_{pq} = n_{\alpha q} n_{\gamma q} \cdot \left(n^2_{\gamma q} - n^2_i \sin^2 i\right)^{-\frac{1}{2}} \;, \quad q = 2,3, \\
&k^+_{sq} = -k^-_{sq} \equiv k_{sq} = \left(n^2_{\gamma q} - n^2_i \sin^2 i\right)^{+\frac{1}{2}} \;, \quad q = 2,3, \\
&\delta^+_{pq} = -\delta^-_{pq} \equiv \delta_{pq} = \left(2\pi n_{\alpha q}/\lambda n_{\gamma q}\right) \cdot \left(n^2_{\gamma q} - n^2_i \sin^2 i\right)^{1/2} \;, \quad q = 2,3, \\
&\delta^+_{sq_0} = -\delta^-_{sq_0} \equiv \delta_{sq} = (2\pi/\lambda) \cdot \left(n^2_{\gamma q} - n^2_i \sin^2 i\right)^{1/2} \;, \quad q = 2,3,
\end{aligned}$$

where the q subscript distinguishes the birefringent plates. Using the above definitions, (14A) and (31A) can be written as

$$\frac{E^+_{m4}}{E^+_{m1}} = C\Big\{\cos(\delta_{m2} d_2)\cos(\delta_{m3} d_3) - K_{m23}\sin(\delta_{m2} d_2)\sin(\delta_{m3} d_3) + j\Big[K_{m12}\sin(\delta_{m2} d_2)\cos(\delta_{m3} d_3) + K_{m34}\cos(\delta_{m2} d_2)\sin(\delta_{m3} d_3)\Big]\Big\}^{-1}, \tag{8}$$

where the subscript m is equal to either p or s and

$$K_{m12} = \frac{1}{2}\left[\frac{k_{m2}}{k_{mi}} + \frac{k_{mi}}{k_{m2}}\right] = \frac{1 + r_{m12}^2}{1 - r_{m12}^2}, \quad m = p, s,$$

$$K_{m23} = \frac{1}{2}\left[\frac{k_{m3}}{k_{m2}} + \frac{k_{m2}}{k_{m3}}\right] = \frac{1 + r_{m23}^2}{1 - r_{m23}^2}, \quad m = p, s,$$

$$K_{m34} = \frac{1}{2}\left[\frac{k_{mi}}{k_{m3}} + \frac{k_{m3}}{k_{mi}}\right] = \frac{1 + r_{m34}^2}{1 - r_{m34}^2}, \quad m = p, s,$$

$$C = \exp\left[j(2\pi n_i \cos i/\lambda)(d_2 + d_3)\right] .$$

2.2 OPTICAL COMPENSATORS

Optical compensators are used to change the polarization state of an electromagnetic wave in a controlable fashion. Most often they are used to convert an elliptically polarized wave into a linearly polarized wave or vice-versa. Eqtn. (8) will be used, eventually, to describe how various types of compensators change the polarization state of the incident wave.

If the incident wave is elliptically polarized, then its p and s electric field components are out of phase with respect to each other. We may characterize this elliptic polarization by

$$E_{p1}^{+} / E_{s1}^{+} = \tan\alpha_i \exp(j\Delta_i) , \tag{9}$$

where Δ_i is the phase of the p component minus the phase of the s component and $\tan\alpha_i$ is the ratio of the absolute value of the p component to the absolute value of the s component. In other words, the right hand side of (9) is the polar representation of the left hand side. Similarly, the elliptic polarization of the transmitted, or output, wave

can be characterized by

$$E_{p4}^{+} / E_{s4}^{+} = \tan \alpha_0 \exp j\Delta_0 . \tag{10}$$

If we define

$$\frac{E_{p4}^{+} / E_{s4}^{+}}{E_{p1}^{+} / E_{s1}^{+}} = \frac{E_{p4}^{+} / E_{p1}^{+}}{E_{s4}^{+} / E_{s1}^{+}} \equiv \tan \alpha \exp j\Delta , \tag{11}$$

then, by using (9) and (10), we find that

$$\tan \alpha_0 = \tan \alpha \tan \alpha_i ; \quad \Delta_0 = \Delta + \Delta_i . \tag{12}$$

Eqtns. (12) describe how the polarization state of the incident wave is effected by the transmission characteristics the two plate geometry. It is important to recognize that Δ and α are functions of the plate thicknesses, the principal indices of refraction of both plates, the refractive index of the surrounding medium, the angle of incidence, and the vacuum wavelength; Δ and α are <u>not</u> functions of Δ_i and α_i . The subsequent discussion shall be confined mainly to the problem of converting elliptically polarized light to linearly polarized light. This is a rather arbitrary restriction but it is not serious and has been imposed for the sake of convenience.

To obtain a linearly polarized output wave we must have $\exp(j\Delta_0) = \pm 1$, that is, the p and s components of the out put electric field must have a phase difference of zero or an integral multiple of π . If the two plate simplified geometry is to function as an optical compensator then some property or parameter of the geometry must be adjustable in order to satisfy

$$\Delta + \Delta_i = m\pi , \quad m = 0, \pm 1, \pm 2, \ldots \tag{13}$$

Before explaining how (13) is achieved by various types of compensators it is useful to consider some analytical expressions for Δ .

In the geometrical optics approximation, all of the two media reflectances in (8) are set equal to zero, then (8) is substituted into (11)

to obtain

$$\tan \alpha_a = 1 \; ; \; \Delta_a = (\delta_{s2} - \delta_{p2})\, d_2 + (\delta_{s3} - \delta_{p3})\, d_3 \; , \tag{14}$$

where the subscript a has been attached to α and Δ to identify them as geometrical optics, or approximate, solutions.

By substituting (8) directly into (11), the resulting expressions for α_e and Δ_e are rather complicated. The subscript e is attached to α and Δ to identify them as exact, or electromagnetic, solutions. An approximate expression for Δ_e can be obtained, however; this is

$$\Delta_e \cong \Delta_a + \Phi \cos\theta \sin \Delta_a \; , \tag{15}$$

where* $$\Phi = (K_{s12} + K_{s34} + K_{p12} + K_{p34} - 4)/4 \; ,$$

$$\theta = (\delta_{s2} + \delta_{p2})\, d_2 + (\delta_{s3} + \delta_{p3})\, d_3 \; .$$

In obtaining (15), several approximations were used. The first is that

$$K_{m23} \cong 1 \quad , \; m = p, s \; . \tag{16}$$

In most optical compensators employing two plates, it happens that the two plates are fabricated from the same crystalline material with the principal dielectric axes of one plate oriented differently from those of the other plate. Since the numerical values of the principal refractive indices of most crystals never differ by large amounts, we might expect that $r_{m23}^2 \cong 0 \; , \; m = p, s$, which, in turn, implies (16). The other approximations used are

$$\left| K_{m12} - K_{m34} \right| \ll K_{m12} + K_{m34} \ll 4 \; , \; m = p, s \; .$$

The important feature of (15) is that the maximum difference between Δ_e and Δ_a is given approximately by Φ .

---

* Note that θ is given a different definition from that used in Chapt. I.

2.2.1 RETARDATION PLATES

Probably the most common of all the various types of compensators is the retardation plate[16,17,18]. This compensator, in its simplest form consists of a single crystalline plate which is placed in the path of the incident wave such that the plate surfaces are perpendicular to the incident wave normal. In this instance, the p (s) component of the transmitted electric field is simply the x (y) component. For retardation plates, Δ_e and α_e are fixed quantities, therefore, (13) is satisfied by rotating the plate about the z axis. This, in effect, changes Δ_i. By considering medium 2 to be identical to medium 3 with principal refractive indices n_α and n_y, we obtain

$$4(\Phi+1) = \frac{n_\alpha}{n_i} + \frac{n_i}{n_\alpha} + \frac{n_y}{n_i} + \frac{n_i}{n_y} \quad . \tag{17}$$

Significantly, Φ does not depend upon the difference between n_α and n_y. If we consider that $n_i = 1$ (air environment) and n_α and n_y both on the order of 1.5, then $\Phi \cong 0.08$ radian $\cong 5^\circ$. By choosing an immersion medium with n_i between n_α and n_y, Φ can be substantially reduced.

If the plate is rotated so that the incident wave is compensated then, because $\pm \tan \psi_e$ is not necessarily equal to unity, the electric field vector of the transmitted wave will not point in the direction predicted by geometrical optics. The implications of this and other effects owing to multiple internal reflections in single plate retarders are contained in App. B. A very general technique for using a retardation plate in conjunction with a rotating analyzer is discussed in App. F.

A variable retardation plate is the Soleil compensator[16]. This device actually can be thought of as two plates with $n_{\alpha 2} = n_{y3}$ and $n_{\alpha 3} = n_{y2}$. One plate consists of two wedges so that its thickness can be varied in a known fashion. By adjusting the thickness of the plate, Δ_e can be varied until (13) is satisfied. For the Soleil compensator, Φ is given by (17), if we attach either a 2 or 3 subscript to n_α and n_y.

2.2.2 ROTARY COMPENSATORS

The phase difference Δ_e , as was mentioned earlier, is a function of the angle of incidence. The operation of rotary compensators is based on this fact. Rotary compensators typically consist of a single birefringent plate or a two plate sandwich. To obtain compensation of the incident elliptically polarized wave, the plate (or plates) is rotated about an axis parallel to the plate surfaces and perpendicular to the incident wave normal, thus changing Δ_e , until (13) is satisfied. The y axis in our geometry (Fig. 1.1) would correspond to the axis of rotation. Multiple reflection effects in the two most common commercially available rotary compensators are discussed in detail in App. C. The use of an index matching environment for eliminating multiple internal reflections is treated in App. D.

A single isotropic slab can also be used as a rotary compensator. More properly one should probably call it a rotary retardation plate since it will not compensate all possible ellipticities of the incident wave. This proposed device, which is treated in App. C, offers interesting possibilities for examining elliptically polarized waves in the far infrared and submillimeter wave regions of the electromagnetic spectrum.

The limited range of the isotropic rotary retardation plate (it will not compensate circularly polarized light) is not particularly serious in ellipsometry experiments. While ellipsometric measuring methods are discussed with greater detail in the following chapter, the general procedure in these methods is to reflect a linearly polarized wave from a specimen surface, then, after examining the elliptically polarized reflected wave, one can use appropriate equations to calculate the optical constants of the specimen. Equations similar to (12) can be derived relating the polarization state of the reflected wave to the polarization state of the the incident wave; when this is done, the ellipticity (minor to major axis

ratio) of the reflected wave is given by*

$$\text{ellipticity} = \tan\left\{\frac{1}{2}\sin^{-1}\left[\frac{2\sin\Delta\tan\alpha\tan\alpha_i}{1+\tan^2\alpha\tan^2\alpha_i}\right]\right\} .$$

The significance of the above equation is that the ellipticity of the reflected wave can always be brought into the measuring range of the isotropic rotary retardation plate by adjusting the direction of the electric vector of the wave incident upon the specimen surface. This adjustment reflects itself in α_i . When the electric field vector of the incident wave is wholly contained in the specimen plane of incidence, $\alpha_i = 90^\circ$ and when it is perpendicular to the plane of incidence, $\alpha_i = 0^\circ$.

---

* This solution for the ellipticity of the reflected light is discussed in greater detail by D. A. Holmes and D. L. Feucht, J. Opt. Soc. Am. 55, 577 (1965). See also Appendix G.

CHAPTER 3

Theory of an Ellipsometric Technique for Investigating Weakly Absorbing Substances

As mentioned in the INTRODUCTION the ultimate aim of this research was to measure the optical constants of a stressed specimen of crystalline GaAs. The main interest is in how the optical constants of a stressed specimen differ from those of an unstressed specimen, therefore, the actual measurements were to be taken, first for an unstressed specimen, and then on the same specimen in a specific state of stress. In this chapter I discuss the theory for the measurement technique as it has so far been developed.

3.1 PRELIMINARY CONSIDERATIONS

Consider the three media geometry depicted in Fig. 3.1. Medium 1 is considered to have unity index of refraction. Medium 2 is a weakly absorbing isotropic* material with complex refractive index** $n_2 = n - jk$ and thickness d. By weakly absorbing we imply that $k \ll n$; in fact, k may be a couple of orders of magnitude less than n. The third medium, or the substrate, is isotropic with complex refractive index $n_3 = N - jK$. The incident wave normal makes an angle i with the positive z axis.

The incident and reflected electric fields are decomposed into p and s components. The convention for the p components is that E_{p1}^{+} and E_{p1}^{-} are considered positive when they point away from the interface between media 1 and 2. We shall define the amplitude reflectances

---

* Later in this chapter medium 2 will be considered to be anisotropic with principal dielectric axes coinciding with the xyz axes in Fig. 3-1.

** The definitions for n and k used in this chapter are not the same as in Chapt. 1 and App. A.

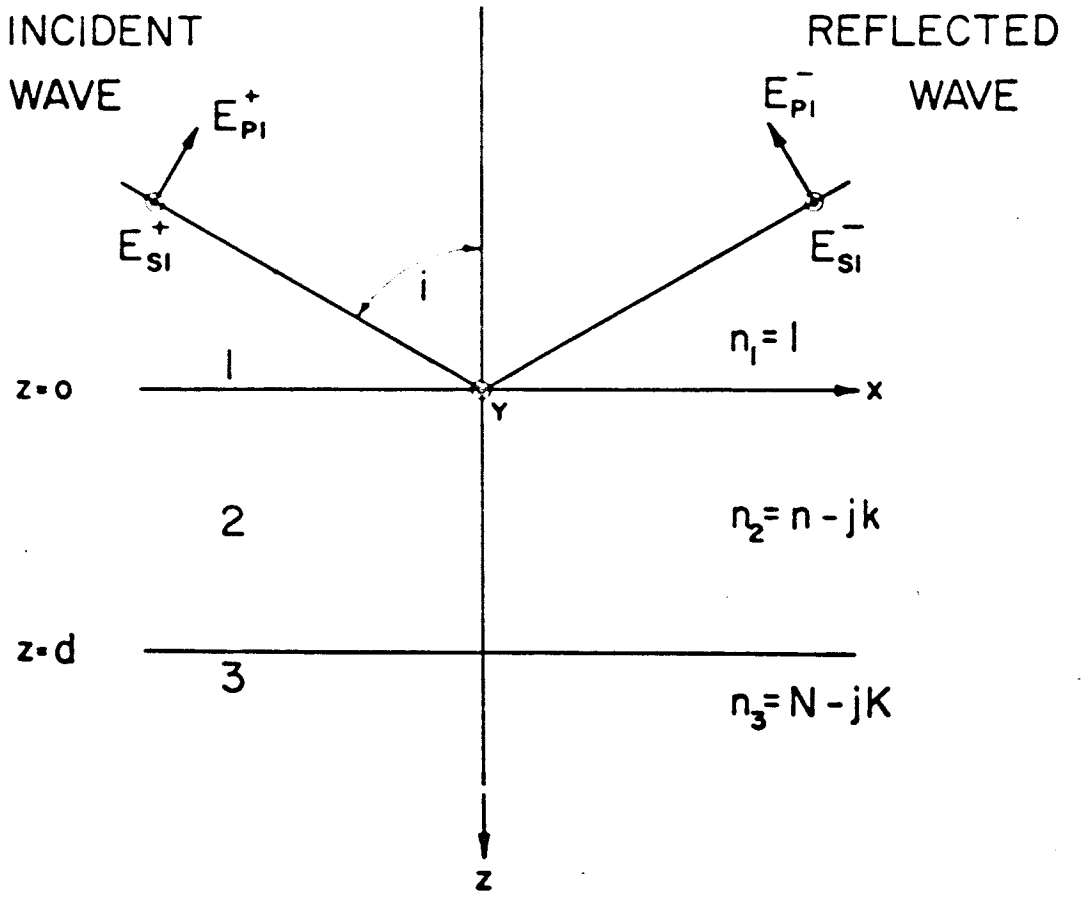


Fig. 3-1 Geometry under consideration. Medium 1 is assumed to be air. Medium 2 is the weakly absorbing material under investigation. Medium 3 is the substrate material. The interfaces between media are assumed to be flat and parallel.

R_p and R_s by

$$R_p \equiv E_{p1}^- / E_{p1}^+ = - E_{x1}^- / E_{x1}^+ \,, \tag{18a}$$

$$R_s \equiv E_{s1}^- / E_{s1}^+ = + E_{y1}^- / E_{y1}^+ \,. \tag{18b}$$

From Chapt. I and App. A, the expressions for R_p and R_s are

$$R_m = \left(r_{m12} + r_{m23} \exp\Gamma\right) / \left(1 + r_{m12} r_{m23} \exp\Gamma\right) \,, \quad m = p, s, \tag{19}$$

where*

$$r_{p12} = \frac{\left(n_2^2 - \sin^2 i\right)^{1/2} - n_2^2 \cos i}{\left(n_2^2 - \sin^2 i\right)^{1/2} + n_2^2 \cos i} \,, \tag{20a}$$

$$r_{p23} = \frac{n_2^2 \left(n_3^2 - \sin^2 i\right)^{1/2} - n_3^2 \left(n_2^2 - \sin^2 i\right)^{1/2}}{n_2^2 \left(n_3^2 - \sin^2 i\right)^{1/2} + n_3^2 \left(n_2^2 - \sin^2 i\right)^{1/2}} \,, \tag{20b}$$

$$r_{s12} = \frac{\cos i - \left(n_2^2 - \sin^2 i\right)^{1/2}}{\cos i + \left(n_2^2 - \sin^2 i\right)^{1/2}} \,, \tag{20c}$$

$$r_{s23} = \frac{\left(n_2^2 - \sin^2 i\right)^{1/2} - \left(n_3^2 - \sin^2 i\right)^{1/2}}{\left(n_2^2 - \sin^2 i\right)^{1/2} + \left(n_3^2 - \sin^2 i\right)^{1/2}} \,, \tag{20d}$$

$$\Gamma = - j \left(4\pi d / \lambda\right) \cdot \left(n_2^2 - \sin^2 i\right)^{1/2} \,.$$

By characterizing the polarization state of the incident and reflected waves analogously to Chapt. 2, we have

$$E_{p1}^+ / E_{s1}^+ \equiv \tan\psi_i \exp j\Delta_i \;\; ; \;\; E_{p1}^- / E_{s1}^- \equiv \tan\psi_r \exp j\Delta_r \,. \tag{21}$$

Using (18) and (21) in

$$R_p / R_s \equiv \tan\psi \exp j\Delta \,, \tag{22}$$

---

* In keeping with the sign convention used to define R_p, the signs of r_{p12} and r_{p23} have been reversed from the convention used in App. A.

we obtain

$$\tan\psi_r = \tan\psi \tan\psi_i \quad ; \quad \Delta_r = \Delta + \Delta_i \tag{23}$$

In an ellipsometry experiment, the incident light is usually linearly polarized, therefore, Δ_i can be set to zero in (23). In this instance ψ_i is determined and hence, measurable, by the setting of a polarizer and, since Δ_r and ψ_r are measurable by a compensator-analyzer combination*, Δ and ψ can be determined through (23). From (19) it is apparent that the ratio R_p/R_s is a function of i, n, k, N, K, and d/λ. We shall consider, that λ is known, that i is measurable, and that the optical constants N and K of the substrate are either known or independently measurable. The fundamental problem then presents itself: After determining i, N, K, ψ_i, ψ_r, Δ_r, and λ, how can one use (22) to accurately calculate n, k, and d? This is the problem to which we address ourselves in the following sections. The determination of k will receive the most attention.

For illustrative purposes, some numerical examples are worked out in later sections. The optical constants of medium 2 are taken to be those corresponding to GaAs crystals of the type commonly used for injection lasers. To satisfy the condition of weak absorption we shall assume that the GaAs crystal is held at a temperature of 77°K and that the crystal is illuminated by radiation from a GaAs laser also held at 77°K. By using a double monochromator, Turner and Reese[19] have performed transmission type absorption measurements at 77°K on samples of GaAs that were doped to concentrations comparable to those present in the n, p, and $p+$ regions of the GaAs injection laser. The absorption coefficient at λ = 8400A (approximately the GaAs laser wavelength) was found to vary from $\alpha = 15\ cm^{-1}$ to $\alpha = 550\ cm^{-1}$. From the relation $\alpha = 4\pi k/\lambda$ we see that $\alpha = 150\ cm^{-1}$ corresponds roughly to $k = 0.001$. Marple[20] has determined the real part of the complex refractive index

---

* See App. F.

of GaAs at 100°K from prism refraction measurements. At 8400A, Marple finds $n = 3.58$ and (apparently) believes that this value of n holds for all laser type GaAs. Hence, in the following sections we shall consider ellipsometric techniques for measuring n and k of medium 2, using the values $n = 3.58$ and $0 \leqq k \leqq 0.007$ for illustrative purposes in the numerical work.

3.2 TWO MEDIA PROBLEM (ISOTROPIC GaAs)

Our interest is mainly in the optical constants of medium 2, namely, n and k, therefore, the thickness d is merely another quantity which must be measured. When d is a controllable parameter, one is tempted to suggest that d be made sufficiently large so that medium 2 appears to be semi-infinite to the incident wave, thus eliminating the necessity for determining d. An additional advantage which results when d is very large is that (19) and (22) simplify considerably.

The quantity Γ has a negative real part when $k > 0$ so that

$$\lim_{d \to \infty} \exp \Gamma = 0 \,. \tag{24}$$

Assuming that d is sufficiently large so that $\exp \Gamma \cong 0$ is a valid approximation, (22) becomes

$$\tan\psi \exp j\Delta = \frac{\sin i \tan i - (n_2^2 - \sin^2 i)^{1/2}}{\sin i \tan i + (n_2^2 - \sin^2 i)^{1/2}} \,. \tag{25}$$

When (25) is inverted, we obtain[21,22]

$$n + jk = \left[\sin^2 i + \sin^2 i \tan^2 i \cdot \frac{\cos^2 2\psi - \sin^2 2\psi \sin^2 \Delta}{(1 + \sin 2\psi \cos \Delta)^2} \right.$$

$$\left. + j \frac{\sin 4\psi \sin \Delta \sin^2 i \tan^2 i}{(1 + \sin 2\psi \cos \Delta)^2} \right]^{1/2} \,. \tag{26}$$

In the experimental situation, ψ_i , ψ_r , Δ_r , and i are measured*, (23) is used to compute Δ and ψ , and then (26) is used to compute n and k .

For weak absorption, (25) can be approximated by

$$\tan\psi \exp j\Delta \cong \frac{\sin i \tan i - (n^2 - \sin^2 i)^{1/2} + jnk(n^2 - \sin^2 i)^{-1/2}}{\sin i \tan i + (n^2 - \sin^2 i)^{1/2}} \qquad (27)$$

Now, if k is very small compared to n , $\tan\psi \exp j\Delta$ will be virtually independent of k for all angles of incidence except near the Brewster angle i_B, defined by $i_B = \tan^{-1}(n)$. At the Brewster angle the real part of (27) vanishes and Δ is very nearly 90° while $\tan\psi$ exhibits a linear dependence with k . These features are illustrated in Figs. 3-2 and 3-3. Fig. 3-2 shows that Δ has a slight linear dependence upon k except for angles of incidence near the Brewster angle. For a given value of k , $\Delta \to 180^\circ$ as i decreases away from i_B and $\Delta \to 0^\circ$ as i increases away from i_B . Recall that when $\Delta = 0^\circ$, 180°, the reflected wave is linearly polarized. Fig. 3-3 shows that $\tan\psi$ is almost independent of k except when i is near i_B . For $i \to i_B$, $\tan\psi$ is extremely small, implying that it is necessary to adjust ψ_i so that $\tan\psi_i$ is extremely large in order to obtain a measurable value of ψ_r . The most significant point made by Figs. 3-2 and 3-3 is that measurements of ψ_i , ψ_r , and Δ_r should be taken for angles of incidence very close to the Brewster angle because only then will Δ and ψ be sensitive to the precise value of k .

In an experiment, one has to expect a certain amount of measurement error. An error analysis of the two media geometry is contained in App. E. The main conclusion drawn from the error analysis is that realistic measurement errors in ψ_i and ψ_r cause intolerable errors in the calculation of k . The factor which completely eliminates the two media

---

* The vacuum wavelength λ does not appear in (26), however, it must be kept in mind that k and n are wavelength dependent.

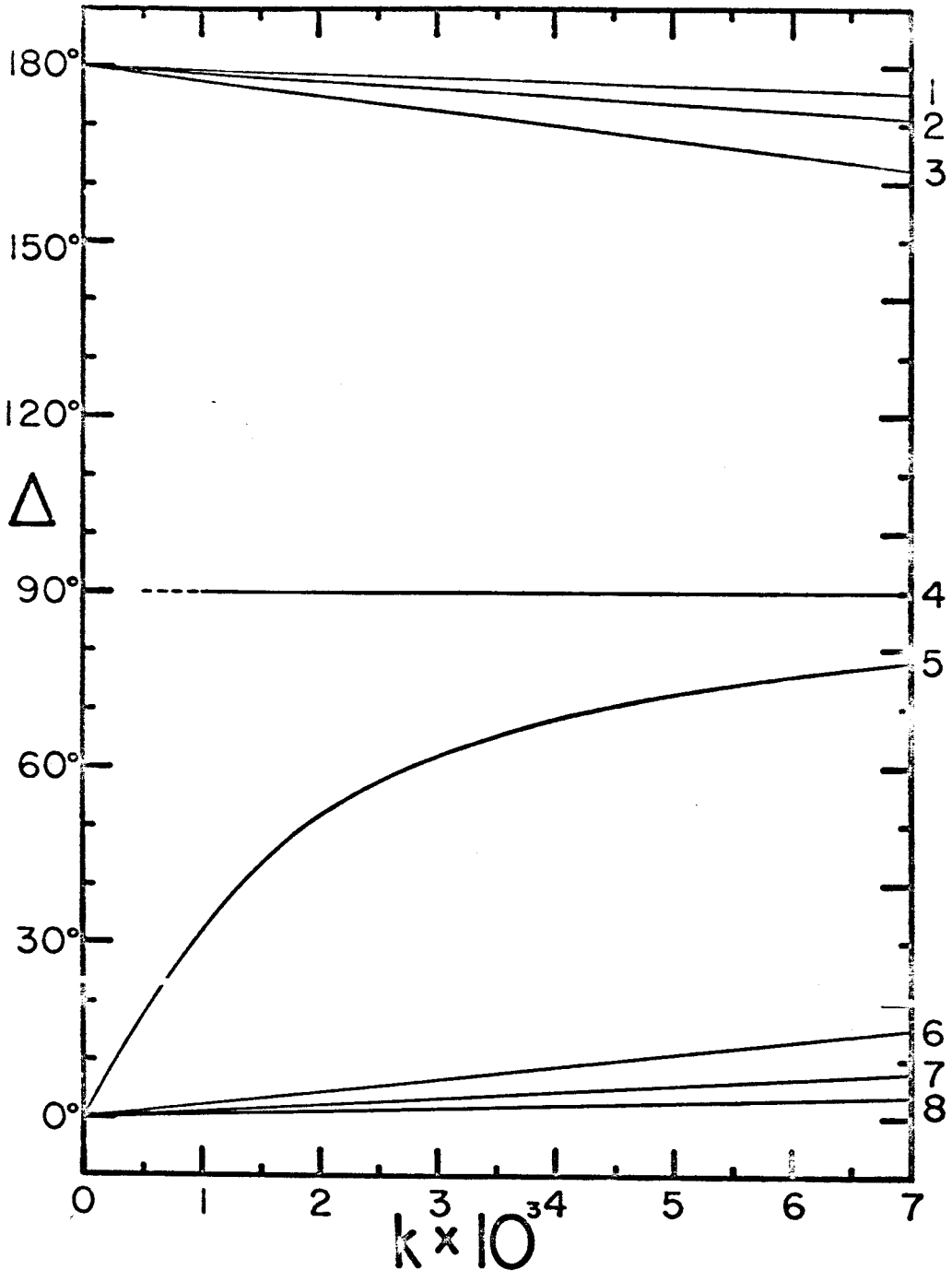


Fig. 3-2 Plot of theoretical values of Δ (degrees, vertical scale), calculated from (25) for the two media geometry, versus $1000k$ (horizontal scale) for some selected angles of incidence, with $n = 3.58$. The numbering specifies the angle of incidence as follows. (1) $i = 74.0^{\circ}$; (2) $i = 74.2^{\circ}$; (3) $i = 74.3^{\circ}$; (4) $i = \arctan(3.58) = 74.39338^{\circ}$; (5) $i = 74.4^{\circ}$; (6) $i = 74.5^{\circ}$; (7) $i = 74.6^{\circ}$; (8) $i = 74.8^{\circ}$.

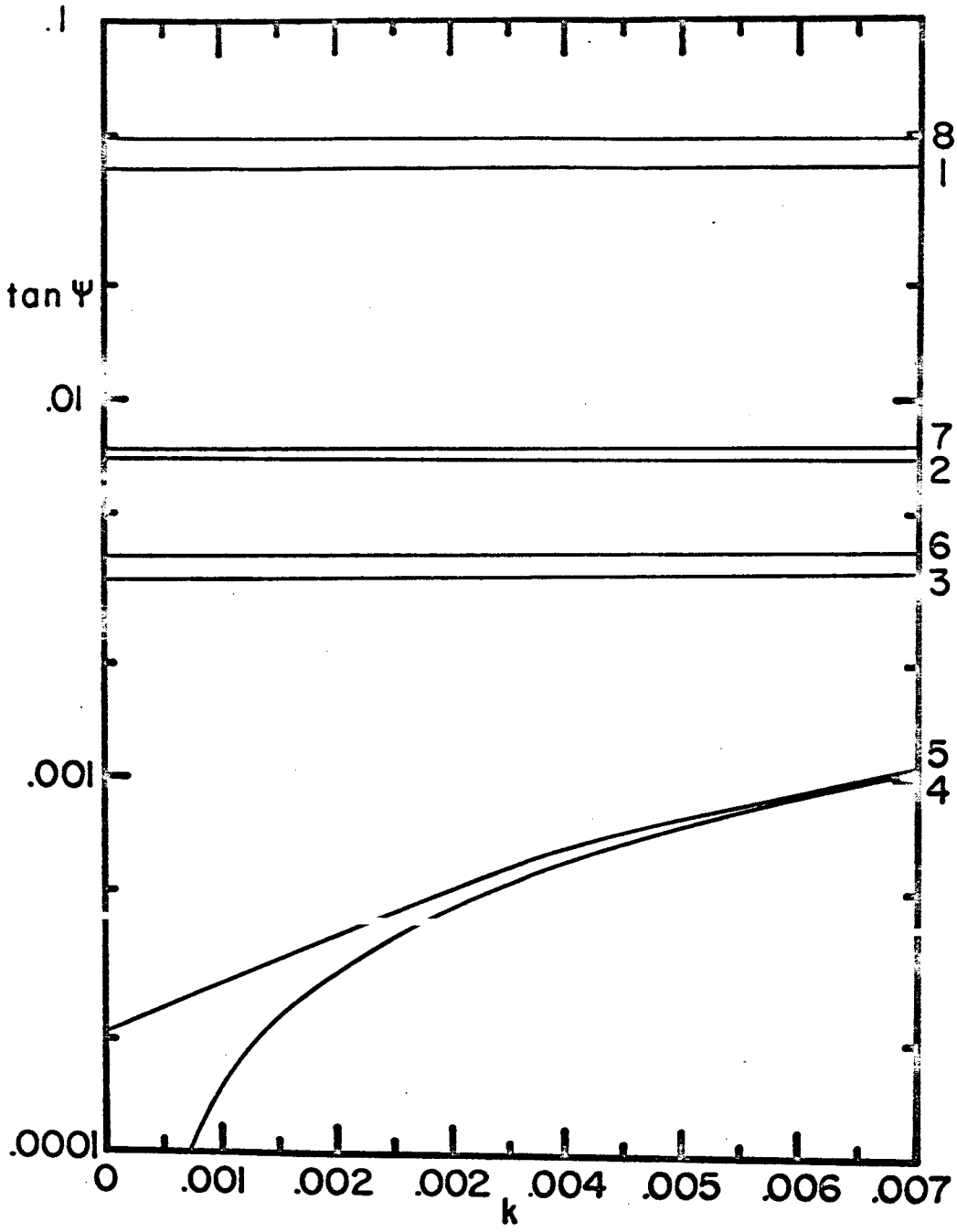


Fig. 3-3 Theoretical values of $\tan\psi$ (vertical axis) for the two media geometry, versus k (horizontal axis) for some selected angles of incidence, with $n = 3.58$. The curves are numbered in accordance with Fig. 3-2.

approach from serious consideration, however, is that one never works with a perfectly collimated incident beam. For example, suppose that the incident beam is slightly divergent and includes angles of incidence between i_B and 74.4° in Fig. 3-2. This is a range in i of less than 0.01°, yet the spread in Δ is greater than 12°, indicating that the reflected beam could not be completely compensated.

3.3 THREE MEDIA PROBLEM (ISOTROPIC GaAs)

The motivation for studying a three media geometry is the possibility that multiple internal reflections can be induced within the slab in such a manner that Δ and Ψ become sensitive functions of n and k while, at the same time, measurement errors do not contribute large errors in the computation of n and k. In addition, we require that Δ and Ψ depend on k in such a way that realistic beam divergences can be tolerated. In order to obtain a reflection from the back surface at $z = d$, we must adjust the value of d so that the approximation $\exp\Gamma \cong 0$ is no longer valid. Physically, this implies that d must be sufficiently small so that reflections from the back surface have a chance to emerge back into medium 1 and significantly effect the polarization state of the reflected wave. When $k \ll n$ (weak absorption), the quantity Γ can be approximated by

$$\Gamma \cong -\frac{4\pi n k d/\lambda}{(n^2 - \sin^2 i)^{1/2}} - j(4\pi d/\lambda)\cdot(n^2 - \sin^2 i)^{1/2} \quad . \tag{28}$$

In order to obtain a significant number of multiple internal reflections, the real part of $-\Gamma$ should not be much larger than unity. In the numerical work in this section we shall use $d = 50\mu$, which, for $k = 0.007$, $n = 3.58$, and $\lambda = 0.84\mu$, makes the real part of $-\Gamma$ approximately 2.5.

In the experimental situation, one is forced to choose a substrate material (medium 3). The substrate material could be air. In this situation, $n_3 = N - jK$ need not be measured since it can be assumed that $N = 1$ and $K = 0$. A possible drawback to using an air substrate, however, is that a certain amount of the energy in the incident wave leaks into medium 3 and is therefore lost. If one is interested in preserving

as much of the incident wave signal strength as possible, then a very high reflectivity substrate can be used. In this section we assume that the latter alternative is desirable and will consider a silver substrate with optical constants[23] $N = 0.2$ and $K = 30.0$.

Another consideration is what angle of incidence to use. Because Δ and ψ both depend upon a number of parameters in addition to i, it is not particularly simple to discern how Δ and ψ will change when i is varied. Some features of the three media geometry can best be illustrated through numerical examples. In Figs. 3-4 and 3-5, Δ and $\tan\psi$ are plotted versus i for some selected values of k. The main point established by the figures is that Δ and ψ can both be sensitive to variations in k over a broad range of i. This is in marked contrast to the two media case where we found Δ and ψ sensitive to the precise value of k only in a very small interval about the Brewster angle. Unfortunately, the figures also show that $\partial\Delta/\partial i$ and $\partial\psi/\partial i$ can be large enough to force stringent requirements on the collimation of the incident beam. Unlike the two media case, however, we are not <u>forced</u> to use an angle of incidence for which beam collimation requirements are unrealistic.

In the three media case the choice of ψ_i (polarizer setting) is not of great importance because a measurable value of ψ_r can nearly always be obtained regardless of the value of ψ_i. It is convenient to choose $\psi_i = \pi/4$ so that a measurement of ψ_r is a direct measurement of ψ, to within experimental error (See eqtn. (23)).

Let us now proceed to demonstrate how one can obtain estimates of n, k, and d/λ (or d if λ is known) through measurements of i, Δ, ψ, N, and K. By using eqtn. (19), eqtn. (22) can be written as a quadratic equation in $\exp(-\Gamma)$, whose solution we can represent by

$$\exp(-\Gamma) = F(n, k, i, \Delta, \psi, N, K) = g + jh \,. \tag{29}$$

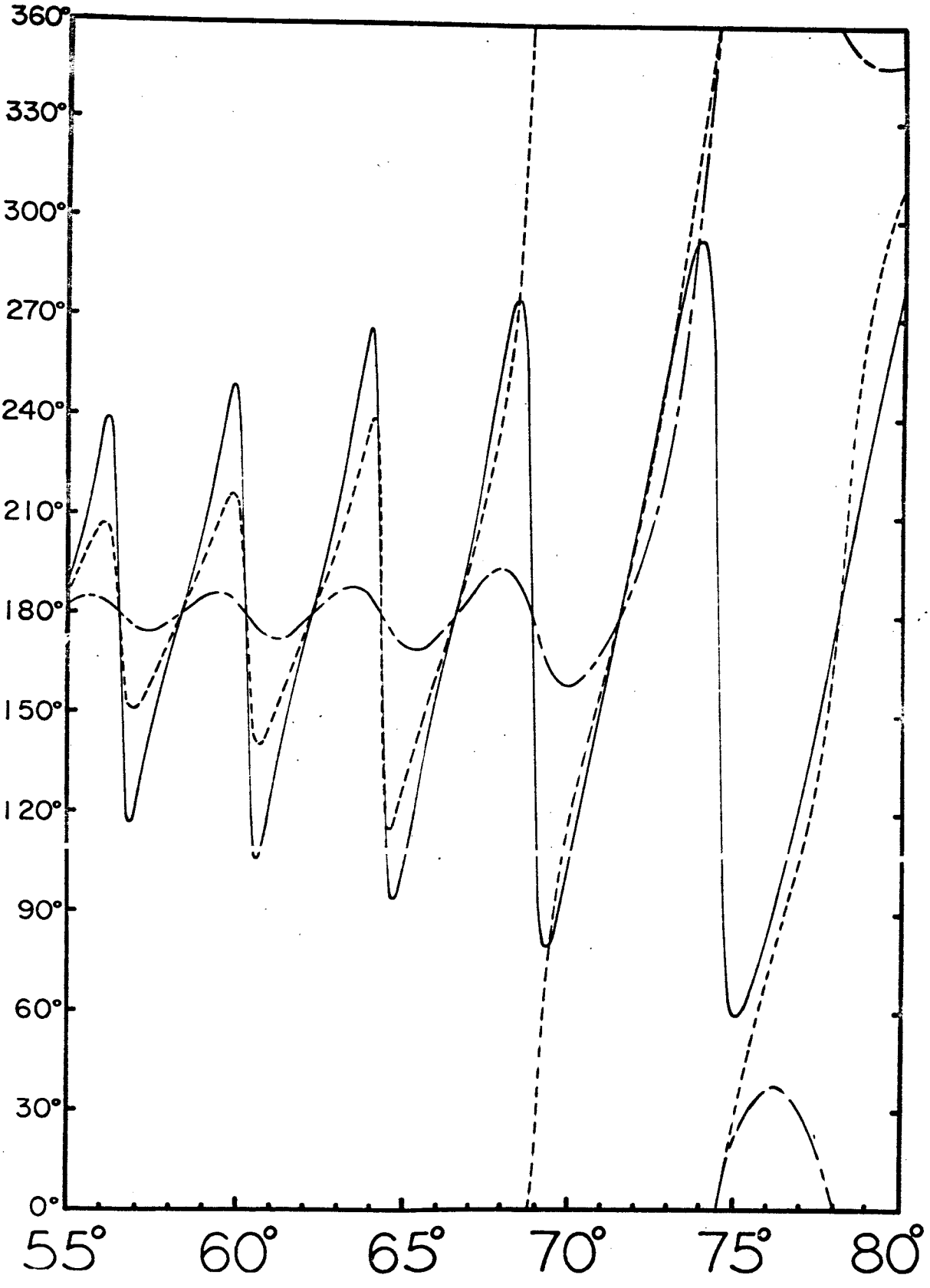


Fig. 3-4 Theoretical plot of Δ (degrees, vertical scale) versus angle of incidence i (degrees, horizontal scale) for the three media geometry with k as a parameter. We have assumed that $n = 3.58$, $d = 50\mu$, $\lambda = 0.84\mu$, $N = 0.2$, and $K = 30.0$.

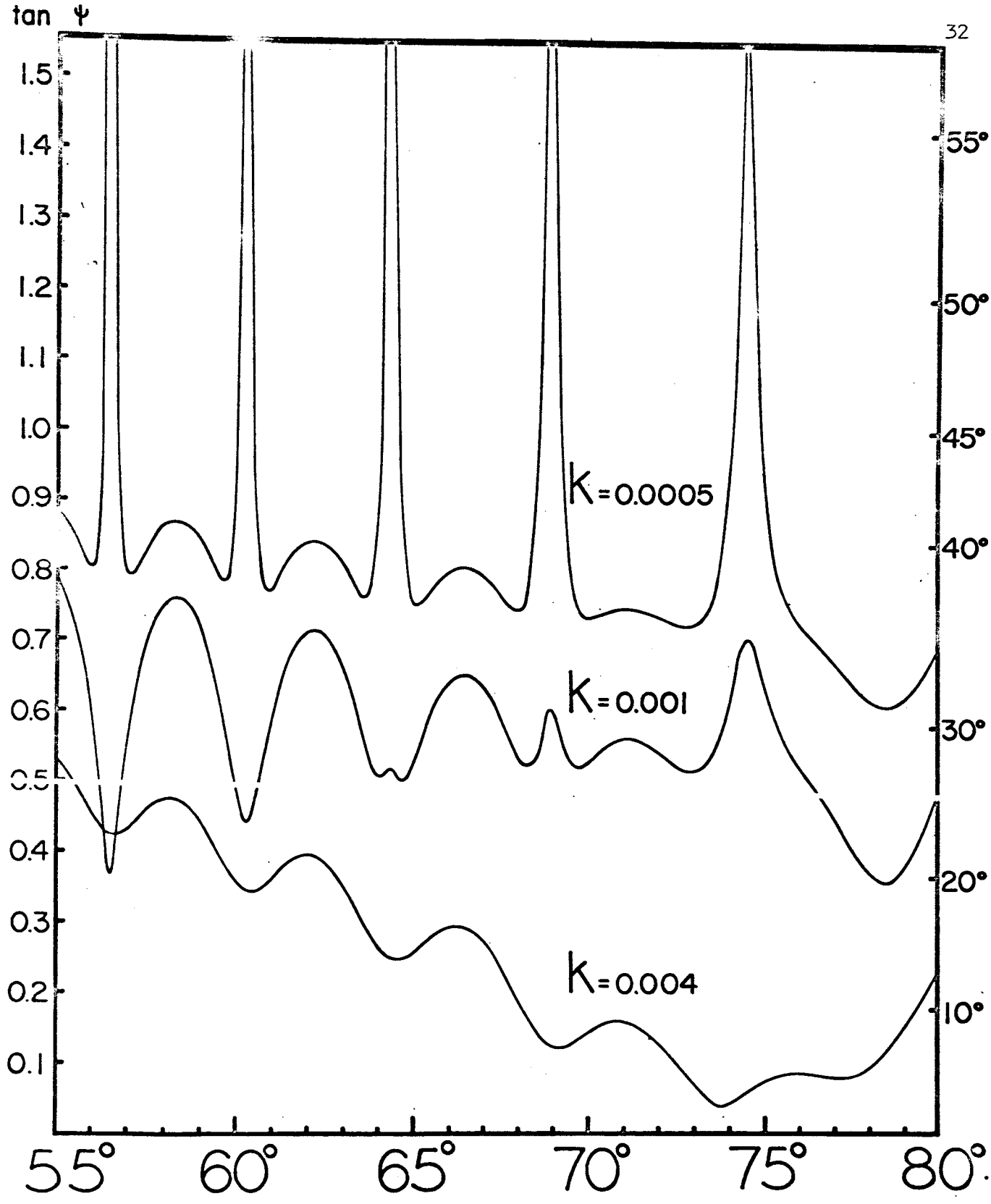


Fig. 3-5 Theoretical plot of $\tan\psi$ (left ordinate) and ψ (right ordinate) versus angle of incidence i in degrees (abscissa) for the three media geometry, with n = 3.58, d/λ = 50/0.84, N = 0.2 and K = 30.0.

Eqtn. (29) contains the three unknowns n , k , and d . We can eliminate d if we consider that measurements have been taken at two angles of incidence, i_1 , and i_2 . The result is

$$(n_2^2 - \sin^2 i_1)^{1/2} / (n_2^2 - \sin^2 i_2)^{1/2} = \ln F_1 / \ln F_2 \quad . \tag{30}$$

Eqtn. (30) is a complex transcendental equation in the complex unknown $n_2 = n - jk$. Reportedly, one can solve this type of equation by an iterative technique on a digital computer[24], however, I have elected to develop an approximate technique.

From our experience with the two media problem, we might suspect that the presence of k in r_{m12} , r_{m23} , $m = p, s$, would have little effect on Δ and Ψ in the three media problem. When Δ and Ψ are calculated by neglecting k everywhere except in Γ , the corresponding approximate values differ only slightly from the exact values; the differences would not be discernable on a plot to the scale of Figs. 3-4 and 3-5. With this approximation, F in (29) is no longer dependent on k . A second approximation for small k , is the expansion (28). Eqtn. (29) may now be written

$$\frac{4\pi n k (d/\lambda)}{(n^2 - \sin^2 i)^{1/2}} = \ln | F(n, i, \Delta, \Psi, N, K) | \quad , \tag{31a}$$

$$4\pi (d/\lambda)(n^2 - \sin^2 i)^{1/2} = 2m\pi + \tan^{-1}(h/g) \quad , \tag{31b}$$

where m is an(unknown) integer. By taking measurements at two angles of incidence i_1 , and i_2 , where $i_1 < i_2$, we can obtain

$$(n^2 - \sin^2 i_1)^{1/2} / (n^2 - \sin^2 i_2)^{1/2} = \ln |F_2| / \ln |F_1| \quad , \tag{32a}$$

$$\frac{4\pi d}{\lambda} = \frac{\tan^{-1}(h_1/g_1) - \tan^{-1}(h_2/g_2)}{(n^2 - \sin^2 i_1)^{1/2} - (n^2 - \sin^2 i_2)^{1/2}} \quad , \tag{32b}$$

$$k = (n^2 - \sin^2 i)^{1/2} \ln |F| / 4\pi (d/\lambda) \quad , \tag{32c}$$

where the subscripts 1 and 2 refer to measurements taken at i_1, and i_2 respectively. Eqtn. (32a) is a real transcendental equation which yields the value of n. After finding n, eqtn. (32b) gives the value of d/λ. Finally, (32c) yields k after n and d/λ have been computed. In the derivation of (32b), a certain amount of care in the section of i_1 and i_2 is assumed. We assume that i_1 and i_2 are sufficiently close together so that the integer m in (31b) is the same for both angles of incidence. The angle $\tan^{-1}(h/g)$ must be assigned to the correct quadrant; if, when this is done, it is found that $\tan^{-1}(h_1/g_1) < \tan^{-1}(h_2/g_2)$, eqtn. (32b) may still be used by adding 2π to the numerator.

A sample calculation, including an error analysis, is presented in App. E. As would be suspected, the sample calculation indicates that the calculated values of n, k, and d/λ do not converge precisely to the true values for "perfect" measurements of i, Δ, ψ, N and K. This is because of the approximations used to derive the solutions (32). There are two very encouraging features of the sample calculation. The first is that gross errors (20-30%) in determining the substrate optical constants resulted in negligible errors in the calculated values of n and k. The second is that no measurement error, considered alone, was drastically amplified in the inversion equations (32). The measurement errors typically contributed less than 10% error in the calculated values of n, k, and d.

The discussion, to this point, has dealt with an ideal geometry and ideal plane waves. In the preparation of plane parallel specimens, one should obviously attempt to simulate the ideal geometry as nearly as possible. A complicating factor which has not been considered here is the possibility that a surface film might exist on the specimen. This would enter analytically as a thin parallel plate specimen between media 1 and 2. The effect of a thin surface film on the ellipsometric determination of optical constants by the two media approach has been recently examined[25].

3.4 THREE MEDIA PROBLEM (ANISOTROPIC GaAs)

Many substances which are normally isotropic, such as GaAs, become anisotropic when stressed. We shall assume that a GaAs slab has been stressed in such a manner that the resultant principal dielectric axes, created by the stress, coincide with the xyz axes in Fig. 3-1. The GaAs (medium 2) then has three principal refractive indices n_{xx}, n_{yy} and n_{zz}, which can be expressed by

$$n_{mm} = (n - jk) + (n_m - jk_m) \quad , \quad m = x, y, z \, . \tag{33}$$

In (33), the quantities n_x, k_x, etc., depend on crystallographic orientation and the state of stress, and reduce to zero when the stress is removed. Note that n_x, k_x, etc., need not necessarily be positive quantities.

The reflected wave is now described by

$$R_m = \left(r_{m12} + r_{m23} \exp\Gamma_m\right) / \left(1 + r_{m12} r_{m23} \exp\Gamma_m\right) \quad , \quad m = p, s \, , \tag{34}$$

where

$$r_{p12} = \frac{n_{xx} n_{zz} \cos i - \left(n_{zz}^2 - \sin^2 i\right)^{1/2}}{n_{xx} n_{zz} \cos i + \left(n_{zz}^2 - \sin^2 i\right)^{1/2}} \, , \tag{35a}$$

$$r_{p23} = \frac{n_3^2 \left(n_{zz}^2 - \sin^2 i\right)^{1/2} - n_{xx} n_{zz} \left(n_3^2 - \sin^2 i\right)^{1/2}}{n_3^2 \left(n_{zz}^2 - \sin^2 i\right)^{1/2} + n_{xx} n_{zz} \left(n_3^2 - \sin^2 i\right)^{1/2}} \, , \tag{35b}$$

$$r_{s12} = \frac{\cos i - \left(n_{yy}^2 - \sin^2 i\right)^{1/2}}{\cos i + \left(n_{yy}^2 - \sin^2 i\right)^{1/2}} \, , \tag{35c}$$

$$r_{s23} = \frac{\left(n_{yy}^2 - \sin^2 i\right)^{1/2} - \left(n_3^2 - \sin^2 i\right)^{1/2}}{\left(n_{yy}^2 - \sin^2 i\right)^{1/2} + \left(n_3^2 - \sin^2 i\right)^{1/2}} \, , \tag{35d}$$

$$\Gamma_p = -j \left(4\pi n_{xx} d / \lambda n_{zz}\right) \left(n_{zz}^2 - \sin^2 i\right)^{1/2} \, , \tag{36a}$$

$$\Gamma_s = -j \left(4\pi d / \lambda\right) \left(n_{yy}^2 - \sin^2 i\right)^{1/2} \, . \tag{36b}$$

The ellipsometric problem is now one of determining the unknowns n_x, k_x, n_y, k_y, n_z, and k_z through measurements of i, Δ, and Ψ, assuming that n, k, N, K, and d/λ are known. This is a rather formidable problem but its solution is facilitated by the fact that the unknowns are adjustable. Consider, for example, that the slab initially is unstressed. Now suppose that a sufficiently small stress is applied so that n_x, n_y, n_z all are several orders of magnitude smaller than n, and k_x, k_y, k_z are all several orders of magnitude smaller than k. In this instance, (35) can be considered, for all practical purposes, to be identical to (20); i.e., the stress induced anisotropy does not measurably affect the two media reflectances. The situation with the $\exp\Gamma_p$ and $\exp\Gamma_s$ factors in (34) can be quite different, however. Expanding Γ_s, for example, we have

$$\Gamma_s \cong -j\frac{4\pi d}{\lambda}\left([n-jk]^2-\sin^2 i\right)^{1/2} - j\frac{4\pi d}{\lambda}\cdot\frac{(n-jk)(n_y-jk_y)}{\left([n-jk]^2-\sin^2 i\right)^{1/2}} \quad . \tag{37}$$

Because d/λ is large (approximately 50/0.84 in this discussion), the second term in (37) can have an imaginary part which would be meaningful to the trigonometric functions in $\exp\Gamma_s$. For example, if $n_y \cong 0.001$, we have $4\pi n_y d/\lambda \cong \pi/4$.

In the following discussion we shall consider the simplest possible anisotropy, namely, that of a uniaxial birefringence for which

$$n_{yy} = (n-jk) + (n_s - jk_s) \quad ,$$

$$n_{xx} = n_{zz} = (n-jk) + (n_p - jk_p) \quad . \tag{38}$$

For this type of anisotropy, the plate appears to be isotropic for both p and s components of incident electric field. Both Γ_p and Γ_s can now be written in the form

$$\Gamma_m \cong \Gamma_u + \gamma_m \quad , \quad m = p, s, \tag{39a}$$

where*, $$\Gamma_u = -j\frac{4\pi d}{\lambda}\cdot\left([n-jk]^2-\sin^2 i\right)^{1/2} \quad ,$$

---

* The subscript u identifies a quantity with the unstressed state.

$$\gamma_m = \frac{-j\frac{4\pi d}{\lambda}\cdot(n-jk)(n_m-jk_m)}{([n-jk]^2-\sin^2 i)^{1/2}}, \quad m = p,s.$$

Since the unknowns n_s, k_s, n_p, and k_p are adjustable, we shall assume that they have been adjusted to allow the approximation

$$\exp\gamma_m \cong 1+\gamma_m, \quad m = p,s. \tag{39b}$$

Using the above approximations, a first order solution is

$$\frac{(\tan\psi \exp j\Delta)-(\tan\psi \exp j\Delta)_u}{(\tan\psi \exp j\Delta)_u} = c_{pu}(n_p-jk_p)-c_{su}(n_s-jk_s), \tag{40}$$

$$c_{mu} = \frac{(-j4\pi d/\lambda)^2(n-jk)(1-r_{m12}^2)(\exp\Gamma_u/\Gamma_u)}{(r_{m12}+r_{m23}\exp\Gamma_u)(1+r_{m12}r_{m23}\exp\Gamma_u)}, \quad m = p,s. \tag{41}$$

Again, measurements must be taken at two angles of incidence in order to obtain two equations of the form (40) which can then be simultaneously solved for the two (complex) unknowns, n_p-jk_p and n_s-jk_s.

Let us suppose that the applied stress is uniaxial. This is possibly the easiest type to apply with an external mechanism. We would suspect, as in stress-strain relations, that n_m and k_m, $m = p,s$, would be directly proportional to the magnitude of the stress up to a certain proportionality limit. If we believe that this proportionality exists* then a convenient check on the validity of the approximations used to obtain (40) is that the calculated values of n_p, k_p, n_s and k_s must all vary linearly with stress.

It is now worthwhile to summarize the conditions and approximations which led to (40). First, the crystal orientation and applied stress must be such that the stressed crystal becomes uniaxially anisotropic with the optic axis pointing in the y direction of the xyz reference frame.

---

* The whole science of photoelastic stress analysis[26] is based upon this belief.

Second, we assume that a measurable change in the polarization state of the reflected wave occurs at such a stress level that the accompanying anisotropy does not alter the two media reflectances from their unstressed values. Third, we assume that the stress-induced anisotropy allows the use of the approximations (39a) and (39b). Fourth, all boundaries remain perfectly flat and parallel with the application of stress. Note here that the change of plate thickness with stress has been ignored. This is probably not justifiable, however, the change in thickness can be accounted for by using stress-strain relations, if the elastic constants of the specimen are known. Fifth, no surface films are present. In closing this paragraph, I draw attention to the fact that no numerical work has been presented. The error analysis for the isotropic case (App. E) should be extended to the anisotropic case.

3.5 STRESS DIRECTION AND CRYSTALLOGRAPHIC ORIENTATION

In this section I discuss the manner in which the GaAs crystal can be oriented and stressed in order to achieve the uniaxial anisotropy considered in Sect. 3.4. It is necessary to first consider the piezobirefringent properties of GaAs. The term "piezobirefringence" was coined by Poindexter[27] as a suggested name for biretringence produced by stress. Poindexter[27] and other workers[28,29] have written articles which review recent research in piezobirefringence. Born and Wolf[30] give a brief introduction to the subject.

Suppose we establish a reference frame within the GaAs crystal such that the orthogonal axes of the reference frame coincide respectively with the three axes of the cubic[31] unit cell. In this reference frame, the electric field $\bar{E}$ is related to the electric flux density $\bar{D}$ by

$$\begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} = [a] \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix}, \quad (42)$$

where $[a]$ is commonly called the index tensor[29]. When the crystal is unstressed, all of the off-diagonal elements of $[a]$ are equal to zero and the diagonal elements are all equal to $a_o = 1/n_o^2$, where $n_o = n - jk$ is the refractive index of the unstressed crystal. If a unidirectional stress V is applied in an arbitrary direction having direction cosines l_V, m_V, n_V with respect to the reference axes, then the elements of the index tensor are given by[29]

$$\begin{bmatrix} a_{11}-a_o \\ a_{22}-a_o \\ a_{33}-a_o \\ a_{12} \\ a_{23} \\ a_{13} \end{bmatrix} = \begin{bmatrix} q_{11} & q_{12} & q_{12} & 0 & 0 & 0 \\ q_{12} & q_{11} & q_{12} & 0 & 0 & 0 \\ q_{12} & q_{12} & q_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & q_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & q_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & q_{44} \end{bmatrix} \begin{bmatrix} l_V l_V V \\ m_V m_V V \\ n_V n_V V \\ l_V m_V V \\ m_V n_V V \\ n_V l_V V \end{bmatrix}. \tag{43}$$

The q_{11}, q_{12}, and q_{44} are called piezo-optic constants. When these constants are known one can then predict how the optical constants will be influenced by an externally applied stress. In Sect. 3.4, I was primarily interested in discussing the ellipsometric measuring technique and therefore did not bring up the subject of piezobirefringence. It is now apparent that the fundamental parameters are the piezo-optic constants, therefore, when the stress direction and crystal orientation are specified so that the uniaxial anisotropy of Sect. 3.4 is created, the ellipsometric measurements and subsequent calculation of n_p, k_p, n_s, and k_s should lead to a determination of q_{11}, q_{12}, and q_{44}.

When $\bar{E}$ and $\bar{D}$ are referred to the principal dielectric axes, which we shall denote by xyz, then

$$\begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = \begin{bmatrix} a_{xx} & 0 & 0 \\ 0 & a_{yy} & 0 \\ 0 & 0 & a_{zz} \end{bmatrix} \begin{bmatrix} D_x \\ D_y \\ D_z \end{bmatrix}. \tag{44}$$

In other words, under a suitable coordinate transformation, namely the transformation from the original crystallographic axes to the principal dielectric axes, the index tensor can be diagonalized.

Now suppose that a wave traveling within the crystal has only an x component of electric field. This could be the case if the unit wave normal vector pointed in either the y or z direction. The direction cosines l, m, and n can be defined such that

$$\begin{aligned} E_1 &= l E_x = a_{xx} l D_x \;;\; D_1 = l D_x , \\ E_2 &= m E_x = a_{xx} m D_x \;;\; D_2 = m D_x , \\ E_3 &= n E_x = a_{xx} n D_x \;;\; D_3 = n D_x . \end{aligned} \tag{45}$$

Substitution of (45) into (42) yields

$$\begin{aligned} (a_{11} - a_{xx}) l + a_{12} m + a_{13} n &= 0 , \\ a_{12} l + (a_{22} - a_{xx}) m + a_{23} n &= 0 , \\ a_{13} l + a_{23} m + (a_{33} - a_{xx}) n &= 0 . \end{aligned} \tag{46}$$

By replacing the x subscripts in (45) with y (or z) subscripts, eqtn. (46) would appear with a_{yy} (or a_{zz}). Therefore, we may write

$$\begin{aligned} (a_{11} - a) l + a_{12} m + a_{13} n &= 0 , \\ a_{12} l + (a_{22} - a) m + a_{23} n &= 0 , \\ a_{13} l + a_{23} m + (a_{33} - a) n &= 0 , \\ l^2 + m^2 + n^2 &= 1 . \end{aligned} \tag{47}$$

Eqtns. (47) are used to diagonalize the index tensor. When (43) is substituted into (47), the diagonal elements of the diagonalized index tensor and the directions of the principal axes can be determined. The determinant of the first three equations in (47) must be equal to zero. When this

determinant is equated to zero, a cubic equation for a results, the solutions of which are a_{xx}, a_{yy}, and a_{zz}. By substituting a_{xx} back into (47) the direction cosines of the x direction can be found. By substituting the solution for a_{yy} (or a_{zz}) back into (47) the direction cosines of the y (or z) directions are determined. We shall now consider an example.

Suppose that the unidirectional stress is applied in one of the $\langle 111 \rangle$ crystallographic directions such that $l_v = m_v = n_v = 1/\sqrt{3}$. When (43) is substituted into (47), the cubic equation for a can be factored into

$$\{a - a_0 - (q_{11} - q_{44} + 2q_{12})(V/3)\}^2 \cdot \{a - a_0 - (q_{11} + 2q_{44} + 2q_{12})(V/3)\} = 0 . \tag{48}$$

By inspecting (48), it is apparent that there are only two distinct solutions for a. Two elements of the diagonalized index tensor are equal, therefore, the unidirectional stress has created a uniaxial anisotropy. We shall choose the roots as follows

$$a_{xx} = a_{zz} = a_0 + (q_{11} - q_{44} + 2q_{12})(V/3) , \tag{49a}$$

$$a_{yy} = a_0 + (q_{11} + 2q_{44} + 2q_{12})(V/3) . \tag{49b}$$

The optic axis of the uniaxially anisotropic crystal is thus taken as the y axis. By substituting (49b) into (47) we find that the y axis, or optic axis, is parallel to the direction of the unidirectional stress. The x and z axes can be chosen in any manner so long as they are perpendicular to each other and to the y axis.

For another example, suppose that the stress is applied in one of the $\langle 100 \rangle$ directions such that $l_v = n_v = 0$ and $m_v = 1$. In this instance (43) can be substituted directly into (42) to yield

$$[a] = \begin{bmatrix} a_0 + q_{12}V & 0 & 0 \\ 0 & a_0 + q_{11}V & 0 \\ 0 & 0 & a_0 + q_{12}V \end{bmatrix} . \tag{50}$$

The index tensor in (50) is already diagonalized to the form in (44). If we choose

$$a_{xx} = a_{zz} = a_o + q_{12}V \,, \tag{51a}$$

$$a_{yy} = a_o + q_{11}V \,, \tag{51b}$$

then the optic axis and the stress direction are both parallel to the y axis.

If we now define the xyz principal axes as those axes corresponding to the geometry of Fig. 3-1, we see that both of the examples considered above will suffice to establish the necessary uniaxial geometry analyzed in Sect. 3.4. The last task is show how q_{11} , q_{12} and q_{44} can be calculated.

The principal elements of the index tensor are given by

$$a_{mm} = 1/n_{mm}^2 \quad , \; m = x, y, z \,, \tag{52}$$

where the n_{mm} are the principal refractive indices. From (38) and the definition $n_o = n - jk$, we have

$$a_{xx} - a_o = a_{zz} - a_o \cong -2\,(n_p - jk_p)/(n-jk)^3 \,, \tag{53a}$$

$$a_{yy} - a_o \cong -2\,(n_s - jk_s)/(n-jk)^3 \,. \tag{53b}$$

When a $\langle 100 \rangle$ direction is parallel to the y axis in Fig. 3-1 we have, using (53) and (51),

$$q_{12} = -2(n_p - jk_p)/(n-jk)^3 \cdot V \,, \tag{54a}$$

$$q_{11} = -2(n_s - jk_s)/(n-jk)^3 \cdot V \,. \tag{54b}$$

When a ⟨111⟩ direction is parallel to the y axis in Fig. 3-1 we have, using (53) and (52),

$$q_{44} = q_{11} + 2q_{12} + \left[6(n_p - jk_p)/(n-jk)^3 \cdot V\right], \tag{55a}$$

$$2q_{44} = -q_{11} - 2q_{12} - \left[6(n_s - jk_s)/(n-jk)^3 \cdot V\right]. \tag{55b}$$

There are two rather subtle points regarding (54) and (55). First, it has been assumed that the absorption mechanisms are governed by crystal symmetry in such a manner that the effective conductivity tensor[6,32] and the dielectric tensor can both be diagonalized with respect to the same set of principal axes. Second, the piezo-optic constants are complex numbers because the crystal is absorbing. Interestingly, all the published works on piezobirefringence deal almost exclusively with non-absorbing crystals.

3.6 MEASUREMENT PROCEDURE

It is worthwhile to consider the experimental steps which lead to a determination of the piezo-optic constants. The GaAs specimen is first mounted in a sample holder with a ⟨100⟩ direction perpendicular to the plane of incidence. The polarization state of the reflected light is measured for the unstressed crystal at a single angle of incidence. Without changing the angle of incidence, a unidirectional stress is applied perpendicular to the plane of incidence until a measurable change in the polarization state of the reflected light is detected and measured. The polarization state of the reflected light is then measured at successively greater stress levels. The process is repeated for at least one more angle of incidence. The zero stress measurements are used to calculate n, k, and d as described in Sect. 3.3. The measurements taken at each stress level are used to calculate $n_p - jk_p$ and $n_s - jk_s$ as described in Sect. 3.4. At each stress level q_{11} and q_{12} are calculated

from (54) in Sect. 3.5. The values calculated for q_{11} and q_{12} should be independent of the particular value of stress below the proportional limit. The measurements described above should then be repeated, in every detail, on the same or a similar specimen with a $\langle 111 \rangle$ direction perpendicular to the plane of incidence. These measurements are used to calculate n , k , d , $n_p - jk_p$, and $n_s - jk_s$. The calculated values for $n_p - jk_p$ and $n_s - jk_s$ and the earlier computed values for q_{11} and q_{12} are then used in (55) to compute q_{44} . Again, the computed values for q_{44} should be independent of the stress level below the proportional limit; furthermore, (55a) and (55b) should yield identical values for q_{44} . As a reproducibility check, many pairs of incidence angles and many stress levels can be used.

The measurement of the quantities Δ and ψ , which define the polarization state of the reflected wave (See eqtn. (21).), can be accomplished by using an optical compensator and an analyzer. While the techniques for this measurement are fairly well known, I have some new thoughts to present on the subject of measuring the polarization state of an electromagnetic wave. These are presented in App. F.

CHAPTER 4

Status of the Ellipsometer

While the previous work (Chapts. 1, 2 and 3) was in progress, efforts were directed toward the development of an ellipsometer. The reason for building an ellipsometer was to create an instrument for investigating the optical properties of stressed laser-like crystals of GaAs for wavelengths in the vicinity of the absorption edge. In the design and construction of the ellipsometer the following general guide-lines were adopted. A pulsed GaAs injection laser held at 77°K was to be used as the light source. The GaAs specimen was to be held at 77°K, or thereabouts, to insure that the source wavelength would be near the specimen's absorption edge. Because the theory for the measuring technique (Chapt. 3) was not completely worked out, it was decided to build an ellipsometer with capabilities for making optical measurements with as much precision as possible.

The mechanical construction of the ellipsometer is virtually complete, however, some calibrating needs to be done before measurements for stressed specimens can be taken. The following sections contain a description of the present state of the ellipsometer. The mechanical design of the ellipsometer is somewhat sophisticated, hence, the following discussion, in places, only highlights some of the basic features. Creative suggestions from Prof. D. Feucht, Mr. E. Litot, and Prof. R. Longini have been invaluable. Most of the detailed mechanical design was done by Mr. Litot. The ellipsometer was built in the Electrical Engineering Machine Shop.

4.1 GENERAL DESCRIPTION

A block diagram (top view) of the ellipsometer is shown in Fig. 4-1. The radiation from a GaAs diode laser source S is converted into a parallel beam by a collimator C. The beam is passed through a polarizer P in order to obtain a linearly polarized beam. The linearly polarized beam is directed

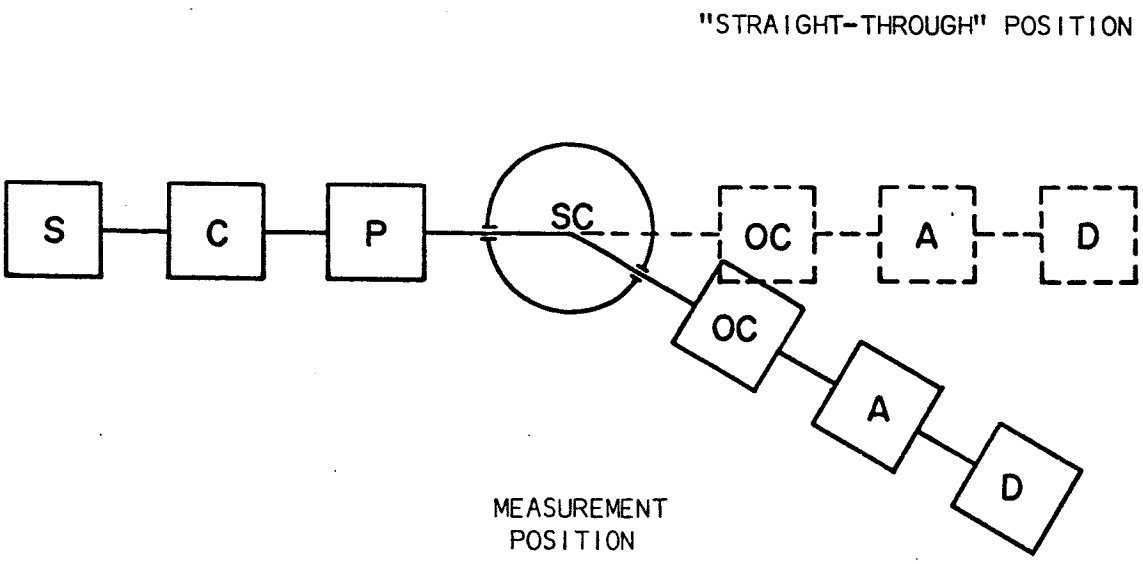


Fig. 4-1 Block diagram of the ellipsometer.

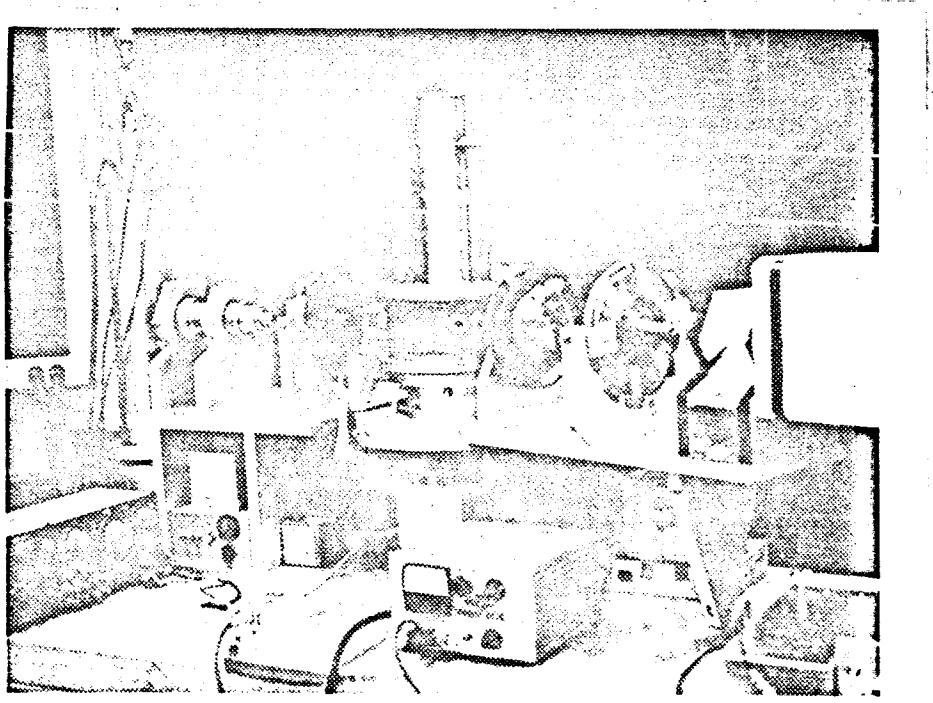


Fig. 4-2 Photograph of the ellipsometer.

upon a GaAs sample in the specimen chamber SC. After reflection from the sample, the polarization state of the reflected beam is measured with an optical compensator OC and an analyzer A. The light emerging from the analyzer is then passed into the detector D.

A photograph of the ellipsometer is shown in Fig. 4-2. The entire ellipsometer assembly is mounted on a 3/4 inch thick aluminum base plate. The surface of the base plate is quite planar; the deviation from perfect flatness is approximately 0.004 inch, or less, over the entire working surface. The base plate rests on the polished surface of a 2-$\frac{1}{4}$ inch thick granite slab which in turn is epoxyed to the top of a sturdy bench. The vertically oriented (brass) cylinder at the left houses a dewar assembly containing the diode laser source. The brass cylinder is mounted on gimbals which permit the cylinder to be tilted from the vertical direction up to a maximum of about 5 deg. The gimbaling arrangement is mounted on a rotary table which allows rotation about a vertical axis and two orthogonal horizontal translations. The rotary table is attached go the base plate. To the right of the source geometry is an elevated platform upon which are mounted the collimator and the polarizer. The collimator is held by two supports and can be "aimed" by adjusting six set screws. The polarizer is mounted in a divided circle which can be "flipped" about a horizontal axis. The sample chamber, compensator, analyzer, and detector are all mounted on another elevated platform which can be termed the detector arm. One end of the arm is fixed to a central pivot while the other end is supported by a "V" leg. The bottom of the "V" leg has a ball-bearing wheel which can roll on the base plate. The essential ingredient of the central pivot is a circular horizontal plate, capable of rotation about a vertical axis, mounted in a supporting structure which is securely fastened to the base plate. The detector arm is bolted to the top of the circular plate. The circular plate is actually a divided circle with the smallest scale divisions being 5 min. of arc. The specimen chamber is mounted on the detector arm so that the

central pivot axis of rotation passes roughly through the specimen itself; thus when the arm is moved, the specimen experiences a rotation with little horizontal translation. The specimen chamber can be given positional adjustments with a gimbal-rotary table combination. The cylinder jutting up from atop the specimen chamber is a stainless steel double-wall dewar which can be filled with coolant; the specimen itself is cooled through a thermal conduction path to a copper heat sink at the bottom of the dewar. The specimen chamber and the volume between the inner and outer walls of the dewar can be evacuated. Vacuum tight optical windows permit the entry and exit of the light beam. To the right of the specimen chamber are two divided circles mounted on the detector arm. The optical compensator is mounted on the circle closest to the specimen chamber while the analyzer is mounted on the other circle. The analyzer divided circle can be "flipped" about a horizontal axis. The last component on the detector arm is a multiplier phototube detector. The white box is a polystyrofoam jacket which can be filled with dry ice to cool the phototube.

Brief descriptions of the ellipsometer components are contained in the following sections.

4.2 SOURCE

A diagram of the source dewar assembly is shown in Fig. 4-3. All of the items shown in the diagram are fixed with respect to each other; their collective motion is achieved by the gimbal-rotary table combination used to adjust the position of the cylindrical brass housing shown in Fig. 4-2. A diode laser in series with a 3 ohm resistor terminates a 3 ohm transmission line. A current pulse travels down the line and the subsequent radiation emitted by the diode passes through an unsilvered portion of the glass dewar. The diode laser is attached to the transmission line so that the emitted radiation direction is roughly horizontal.

The test tube contains liquid oxygen while the dewar contains liquid nitrogen; the reasons for this arrangement can be explained through consideration of the following. When high current densities pass through

the diode, heat is generated because of nonradiative recombination and joule heating[40,*]. If, in this situation, the coolant surrounding the diode is at its boiling temperature, then bubbles will form at the surface of the diode. These bubbles act like spherical lenses and interact with the diode beam in an undesired fashion. In the geometry of Fig. 4-3, the oxygen is cooled to the nitrogen temperature and it was found experimentally that no bubbles were formed in the oxygen. The nitrogen does boil; however, the bubbles are sufficiently far away from the diode so that no single bubble can intercept the entire diode beam. The liquid oxygen is formed by condensation from relatively pure oxygen gas, thus, the liquid oxygen surrounding the diode is free from contamination. Also, the gaseous oxygen flush prevents ice from forming on the diode or test tube walls.

4.3 COLLIMATOR

The collimator, which converts the diverging diode radiation into a parallel beam, is shown schematically in Fig. 4-4 . The laser diode is placed at the position PF. The diode radiation is collected by lens L1. The beam emerging from L1 is focussed by another lens L2 through an aperture A1. The diverging light from the aperture A1 is converted into a parallel beam by the lens L3 and the final aperture A2 is used to control the diameter of the working beam. The lenses L1 and L2 are identical and have 83 mm focal length and 51 mm diameter. The lens L3 has 17 mm focal length and 17.5 mm diameter. The aperture A1 is a pinhole having a 0.01 inch diameter. The aperture A2 is a diaphragm whose circular opening can be varied from 1/64 to 9/64 inch diameter. The distances d1 and d2 are adjustable. Under operating conditions, the diode is positioned at the primary focal point of L1 and the pinhole A1 is positioned at the secondary focal point of L2 and the primary focal point of L3. Experimental testing of the collimator shows that the divergence of the working beam can be reduced to approximately 7 min. of arc and possibly less.

---

* Ref. 40 is a comprehensive review paper on injection lasers, covering nearly all of the published works through late 1963.

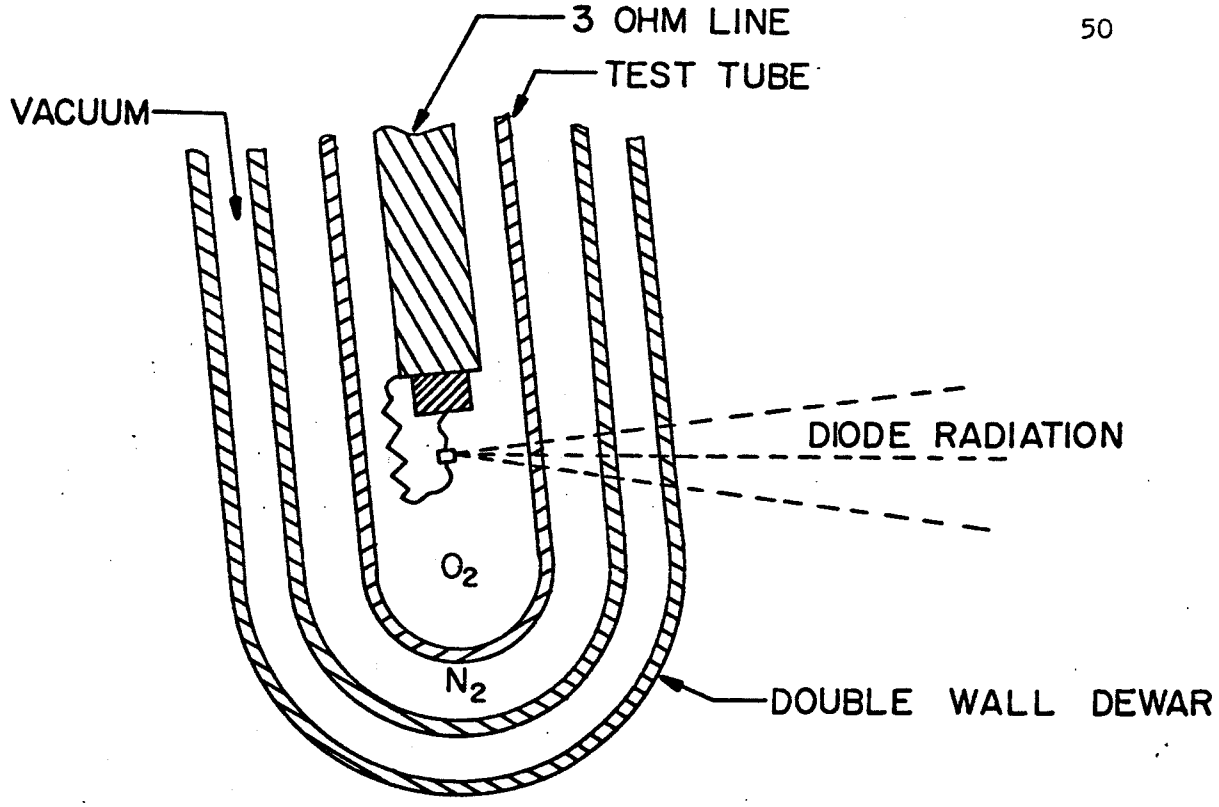


FIG. 4-3. GaAs LASER SOURCE AND DEWAR ASSEMBLY.

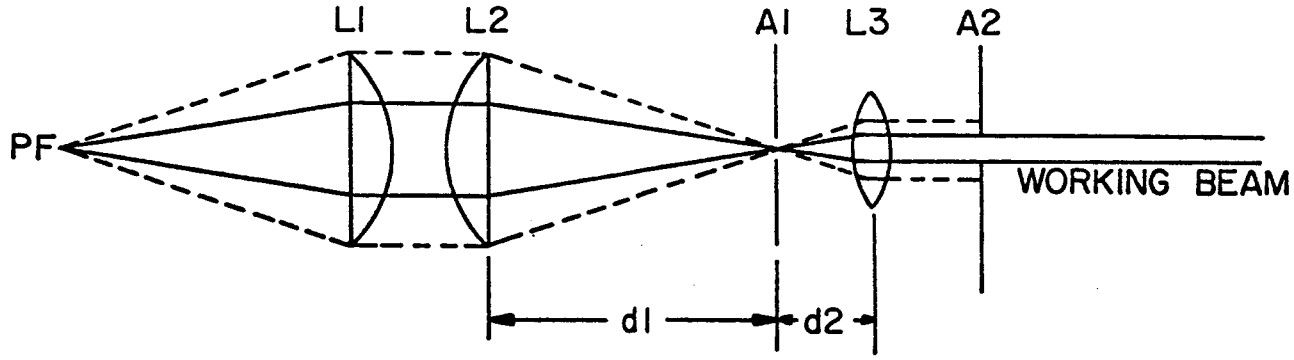


FIG. 4-4. SCHEMATIC DIAGRAM OF COLLIMATOR.

4.4 POLARIZER AND ANALYZER

Both the polarizer and the analyzer are Glan-Thompson prism polarizing devices[8,*]. Relevant information for the particular prisms used in the ellipsometer is as follows. The prism material is single crystal calcite and the physical dimensions of each polarizing device are 8 mm by 8 mm by 20 mm. (The light enters and leaves the square ends.) The optic axis is aligned, within 5 min. of arc, perpendicular to two of the prism sides and parallel to the cemented interface. The cement used is butyl-methacrylate, having a refractive index of 1.51 at 5893A and a useable transmission from 3500A to 23000A. The electric field transmission axis is parallel to the optic axis direction. The Glan-Thompsons are used at normal incidence, however, angles of incidence up to 17 deg. can be tolerated without degradation of performance; thus, the requirements on beam collimation are not strigent. The extinction ratio, as measured in the ellipsometer, is greater than 1,000,000:1. The Glan-Thompsons are mounted on divided circles which can be read by vernier to 1 min. of arc.

Since the calibration of polarizers and analyzers is rarely discussed in the ellipsometry literature we shall give a brief description of our calibration procedure. In ellipsometry it is necessary to know the direction of the electric field transmission axis with respect to the specimen plane of incidence for both the polarizer and the analyzer. In our ellipsometer, the plane of incidence is roughly parallel to the surface of the base plate.

Consider Fig. 4-5. In (a) is shown a plane wave being reflected from an absorbing surface. The incident wave is assumed to be linearly polarized after passing through a polarizer. The electric field amplitude of the incident wave is E_i . The direction of the polarizer transmission axis P-P is located by the angle p . When $p = 0$, E_i is perpendicular to the plane of incidence. After reflection, the incident wave is generally elliptically polarized. The reflected wave is passed through an analyzer whose transmission axis A-A is located by the angle a shown in (c). We shall define E_o as the electric field amplitude of

* Purchased from Crystal Optics, 3959 N. Lincoln Ave., Chicago 13, Ill.

the wave emerging from the analyzer. The components of the reflected wave are given by

$$E_{rp} = R_p E_{ip} = r_p \exp j\delta_p \sin p \; E_i \;, \tag{56a}$$

$$E_{rs} = R_s E_{is} = r_s \exp j\delta_s \cos p \; E_i \;. \tag{56b}$$

After the reflected wave passes through the analyzer we have

$$E_o = E_{rs} \cos a + E_{rp} \sin a \;. \tag{57}$$

By substituting (56) into (57) we have

$$|E_o / E_i|^2 = I_o / I_i = (r_s \cos a \cos p)^2 + (r_p \sin a \sin p)^2 + 2 r_p r_s \cos a \sin a \cos p \sin p \cos \Delta \;, \tag{58}$$

where $\Delta = \delta_p - \delta_s$. Note that I_o is the intensity of the light transmitted through the analyzer. Now suppose that the polarizer and analyzer are "crossed," i.e., $p = a \pm \pi/2$. Eqtn. (58) becomes

$$I_{oc} = (\sin^2(2a)/2)(r_p^2 + r_s^2 - 2 r_p r_s \cos \Delta) \cdot I_i \;. \tag{59}$$

The quantity I_{oc} is a symmetric function of a , that is, $I_{oc}(a) = I_{oc}(-a)$. This symmetry is used in the calibration procedure as follows. The analyzer transmission axis is set at some unknown angle* $a > 0$. The analyzer divided circle will read A_1 . The polarizer is set at ** $a \pm \pi/2$. The polarizer divided circle will read P_1 . For this situation a certain signal will be registered on the detector. The analyzer is rotated to a negative angle such that, with the polarizer crossed, the same signal

---

* Assuming that $-\pi/2 \leq a \leq +\pi/2$, it is possible to determine whether a is greater or less than zero by visually inspecting the Glan-Thompson prism.

** The scale readings which correspond to crossed polarizer and analyzer can be determined by measurements taken with the ellipsometer in the "straight-through" position.

is recorded by the detector. Now the divided circles read A_2 and P_2. The reading on the analyzer divided circle which corresponds to $a = 0$ is then given by $(A_1 + A_2)/2$. The reading on the polarizer scale which corresponds to $p = \pi/2$ is given by $(P_1 + P_2)/2$. By using an evaporated silver film deposited on a microscope slide as a reflector, it has been possible to locate the polarizer and analyzer transmission axes (with respect to the plane of incidence) with a precision of $\pm$ 2 min. of arc.

4.5 SPECIMEN CHAMBER

The chamber containing the GaAs specimen serves several purposes. The specimen can be cooled to approximately 77°K by a thermal conduction path to a heat sink in contact with liquid nitrogen. To prevent convective heat losses and to prevent contaminants from condensing on the specimen surface, the environment surrounding the heat sink and the specimen can be evacuated. Optical windows mounted on the chamber walls allow the working light beam to enter and leave the chamber. The windows are high quality, strain free, optical glass* with a surface flatness of $\lambda/4$ and surface parallelism of 30 sec. of arc. There are six optical windows which are used in pairs. One pair permits the light beam to pass undeviated through the chamber when the specimen is withdrawn. (This is called the "straight-through" position of the ellipsometer.) Another pair of windows permit specimen investigation at angles of incidence between 53.5 and 56.5 deg. The position of each window can be adjusted so that its surfaces are perpendicular to the beam passing through it.

The manner in which stress is applied to the specimen can be understood by examining Fig. 4-6. The specimen is attached to a metal bar which is clamped at both ends. Two equal forces are applied to the bar

---

* Fabricated by Unerti Optical Co., Pittsburgh.

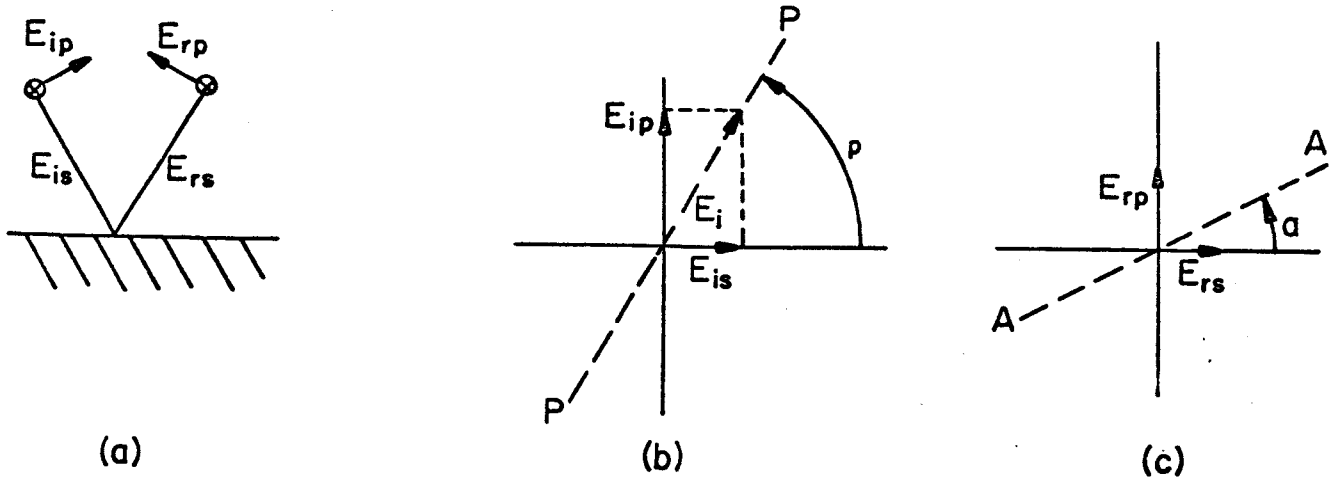


Fig. 4-5 Diagram to aid in describing calibration procedure for polarizer and analyzer.

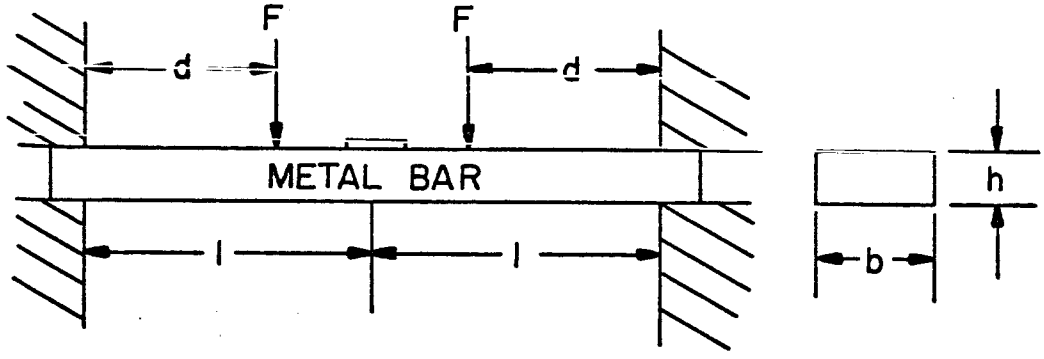


Fig. 4-6 Scheme for applying stress to specimen.

as shown in the figure. When the forces are applied, the top surface of the bar is put in compression. This compressive strain is transmitted to the specimen through the adhesive which bonds the specimen to the bar. The magnitude F of the forces can be controlled by a mechanism which is external to the specimen chamber. The specimen is bonded to the bar at room temperature or higher, therefore, it is desirable that the thermal expansion coefficient of GaAs match that of the metal bar in order to avoid temperature-induced strains in the specimen when it is cooled to nitrogen temperature. The expansion coefficient of GaAs at room temperature[37] is $5.7 \times 10^{-6}/^{o}C$. In addition to having the appropriate expansion coefficient, the metal bar must be sufficiently elastic to introduce, without permanent deformation of the bar, a strain in the specimen large enough to change its optical properties in a measurable way. Sturge[37] applied a strain of roughly 4×10^{-4} to a GaAs sample and was able to detect a significant change in the absorption coefficient in a transmission type measurement. Zirconium was chosen as the metal for the bar. The thermal expansion coefficient of polycrystalline Zr is given as[38] $5.89 \times 10^{-6}/^{o}C$ at a temperature of $20^{o}C$. For given forces F, the amount of strain which can be introduced at the surface of the bar depends on the dimensions, b, d, h, and ℓ (see Fig. 4-6) and on the modulus of elasticity of the bar. An order of magnitude calculation of the strain follows.

When the bar is loaded by the forces F, the deflection of the bar and the strain at its surfaces can be found by using the double integration method for statically indeterminate beams.[39] We shall assume that the GaAs sample presents negligible resistance to the bending of the bar and shall ignore the bar's weight. The bending moment at any cross-section of the bar located between the forces is constant because of the symmetry of the geometry, therefore, the strain on the bar's surfaces is uniform between the forces. This strain is given by

$$\epsilon = (3\ell F/Ebh^2)\cdot(d/\ell)^2 \,, \qquad (60)$$

where E is the bar's modulus of elasticity which, for Zr, is given as approximately 10^7 psi[38]. The quantity ϵ is the compressive strain at all points of the bar's top surface located between the forces. The bars used in the sample chamber have approximately ℓ = 1.5 in., d = 1.0 in., h = 1/8 in. and b = 1/4 in. which yields

$$\epsilon \cong 5 \times 10^{-5} F , \tag{61}$$

where F has the units lbs. The maximum deflection of the bar occurs at its middle and is given by

$$y_{max} = \frac{3F}{bE} \cdot \left(\frac{\ell}{h}\right)^3 \cdot \left(\frac{d}{\ell}\right)^2 \cdot \left(1 - \frac{2d}{3\ell}\right) . \tag{62}$$

For the Zr bar under consideration, (62) becomes

$$y_{max} \cong 0.5 F \tag{63}$$

where y_{max} is in mils (0.001 in.) and F is in lbs. From these rough computations it appears that a Zr bar of dimensions considered above should be satisfactory provided one can obtain good strain coupling from the bar to the GaAs specimen.

At present there remain some unfinished tasks with regard to the specimen chamber. The strain-producing mechanism has to be calibrated in order to establish a correlation between an external adjustment and the magnitude of the strain produced in the bar by the adjustment. The first GaAs specimen (about 60u thick) was bonded to the Zr bar with In solder*. The efficiency of the In for transferring strain from the bar to the specimen has not yet been established. The optical constants of the In used should be measured by the two media approach described in Sect. 3.2. A temperature calibration is necessary to determine the effectiveness of the thermal conduction path for cooling the specimen to (roughly) the temperature of the coolant. Thermocouple feed-throughs have been built into the chamber walls for the latter purpose.

---

* The indium also serves as the high reflectivity substrate considered in Sect. 3.3.

4.6 OPTICAL COMPENSATOR

A mica retardation plate was purchased from Crystal Optics. The mica plate has a phase difference of approximately 90 deg. at 8400 A. The diameter of the plate is 22 mm. Mica cleaves easily into plane-parallel plates uniformity of thickness[36,*] better than 30A. The mica plate is cemented between two glass disks. The disks, which are 2.5 mm thick, have the following properties: surface flatness - $\lambda/4$; surface parallelism - 1 min. of arc; refractive index -1.52. The cement has a refractive index of 1.51. To reduce the effects owing to multiple reflections, the external faces of the disks are coated with a quarter-wave film of magnesium flouride, which has a refractive index of roughly 1.38. In App. F it is shown that multiple reflections need not obstruct the use of a retardation plate, however, multiple reflections do place restrictions on the collimation of the light beam, as shown in App. B. The fast and slow axes of the plate were located (roughly) by the supplier and are marked on the edge of the plate by red and blue dots, respectively. The mica plate is mounted on a divided circle which can be read by vernier to 1 min. of arc.

As shown in App. F it is necessary to know where the fast axis is located with respect to the plane of incidence. Once the polarizer and analyzer are calibrated it is sufficient to establish the compensator divided circle reading which corresponds to the case when the fast axis is parallel to the analyzer transmission axis. Suppose that the ellipsometer is set in the "straight-through" position with crossed polarizer and analyzer as shown in Fig. 4-7. The compensator is placed between the polarizer and the analyzer. The electric field emerging from the polarizer is E_i which has components E_{is} and E_{if}. The electric field emerging from the compensator has components E_{os} and E_{of} where

---

* It is not known whether our retardation plate has a surface uniformity of 30A, however, Crystal Optics has an outstanding reputation for making very high quality optical components.

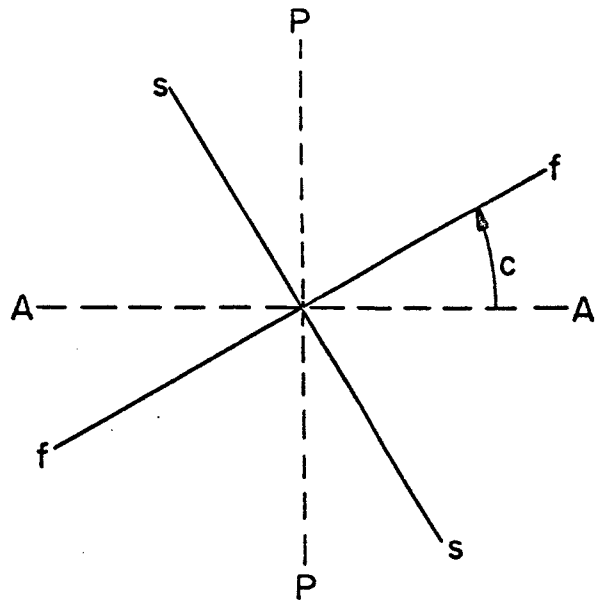


Fig. 4-7 A-A is the analyzer transmission axis.
P-P is the polarizer transmission axis.
f-f is the fast axis of the compensator.
s-s is the slow axis of the compensator.

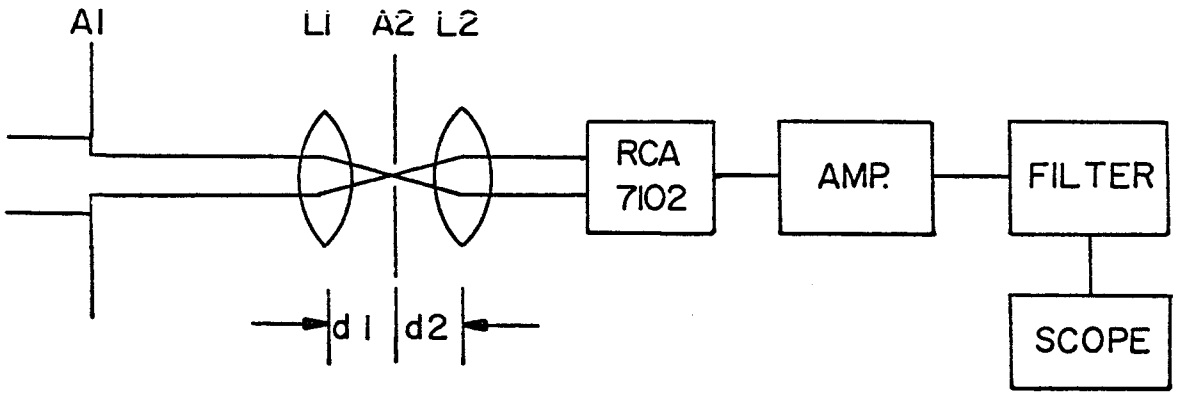


Fig. 4-8 Detection apparatus.

$$E_{os} = T_s E_{is} = t_s \exp j\delta_s \cos c\, E_i \,, \tag{64a}$$

$$E_{of} = T_f E_{if} = t_f \exp j\delta_f \sin c\, E_i \,. \tag{64b}$$

The electric field emerging from the analyzer is E_o where

$$E_o = E_{of} \cos c - E_{os} \sin c \,. \tag{65}$$

The transmitted intensity is found by substituting (64) into (65) to obtain

$$I_o = (\sin^2(2c)/2)\cdot(t_f^2 + t_s^2 - 2t_f t_s \cos[\delta_f - \delta_s])\cdot I_i \tag{66}$$

By setting the fast axis at some unknown angle $c>0$ a certain signal is registered on the detector. The compensator divided circle reads C_1 The fast axis is rotated to angle $-c$ (with a scale reading of C_2) for which the same signal is detected. Then the fast compensator axis is parallel to the transmission axis of the analyzer when the compensator divided circle is set at $(C_1 + C_2)/2$.

4.7 DETECTOR

A diagram of the detector apparatus is shown in Fig. 4-8. Part of the beam reflected from the sample is passed through an aperture A1, focussed through a pinhole A2 by lens L1, recollimated by a lens L2, and detected by an RCA 7102 photomultiplyer tube (S-1 response). Aperture A1 is variable from 1/64 in. dia. to 9/64 in. dia. while A2 is fixed at 0.01 in. dia. L1 and L2 are identical, with focal length 17 mm and dia. 17.5 mm. The distances d1 and d2 are adjustable. The apertures allow resolution of the angle of incidence and insure that the detected beam always strikes the same area of the 7102 photocathode surface. AMP. is a transistor emitter follower-amplifier with a voltage gain of approximately 800 when terminated in 50 ohm. FILTER is an electrical

filter (50 ohm input and output) with a passband from 0 to 1.7 mcs. This passband allows transmission of most of the signal pulse(about 0.6 usec duration) but blocks much of the 7102 dark current noise. The signal pulse from the filter is displayed on an oscilloscope.

APPENDIX A

Birefringent Plate Boundary Value Problem

In Chapt. 2 it will be necessary to know how a wave is transmitted through two plates, therefore, in this appendix we shall consider a four media geometry rather than the three media geometry of Fig. 1-1. In the four media case, the second medium (first plate) is of thickness d_2, the third medium (second plate) is of thickness d_3, and the fourth medium extends from $z = d_2 + d_3$ to $z = \infty$. All media are birefringent with one principal axis in each medium perpendicular to the plane of incidence, or the xz plane.

The "positively" traveling waves in each medium have electric fields which are of the form

$$\bar{E}_q^+ = \bar{E}_q^+ \exp\{j\omega t - j(2\pi n_q^+/\lambda)\hat{s}_q^+ \cdot \bar{r}\}, \tag{1A}$$

while for the "negatively" traveling waves we have

$$\bar{E}_q^- = \bar{E}_q^- \exp\{j\omega t - j(2\pi n_q^-/\lambda)\hat{s}_q^- \cdot \bar{r}\}, \tag{2A}$$

where $q = 1,2,3,4$ for the "positively" running waves and $q = 1,2,3$ for the "negatively" running waves. The magnetic field vectors are found by substituting (1A) and (2A) into the Maxwell equations. The procedure for finding the reflectance and transmittance formulae is conventional in that traveling wave solutions are assumed for the magnetic and electric fields in each medium; the desired formulae are then found by solving the set of equations obtained by requiring the total tangential fields to be continuous at the boundaries $z = 0$, $z = d_2$, and $z = d_2 + d_3$.

Consider first the waves of the extraordinary type which have only x and z components of electric field and y components of magnetic field. In this case the x components of electric field and the y components of magnetic field must be continuous at each boundary. The boundary conditions, when applied to the electric fields, yield

$$E_1^+ + E_1^- = E_2^+ + E_2^- \,, \tag{3A}$$

$$E_2^+ \exp(-j\delta_2^+ d_2) + E_2^- \exp(-j\delta_2^- d_2) = E_3^+ \exp(-j\delta_3^+ d_2) + E_3^- \exp(-j\delta_3^- d_2) \,, \tag{4A}$$

$$E_3^+ \exp(-j\delta_3^+ (d_2+d_3)) + E_3^- \exp(-j\delta_3^- (d_2+d_3)) = E_4^+ \exp(-j\delta_4^+ (d_2+d_3)) \,, \tag{5A}$$

where the x subscript is understood to be affixed to each electric field component, e.g., E_{x1}^+ . The subscripted δ quantities are all of the form

$$\delta_2^+ = 2\pi n_2^+ \cos\theta_2^+ / \lambda \,. \tag{6A}$$

All of the magnetic fields have the form (MKS units)

$$-j\omega\mu_0 \mathcal{H}_{y2}^+ = \partial \mathcal{E}_{x2}^+ / \partial z - \partial \mathcal{E}_{z2}^+ / \partial x \,. \tag{7A}$$

By substituting (1A) into (7A) we have

$$(\omega\mu_0 \lambda / 2\pi) \mathcal{H}_{y2}^+ = k_2^+ \mathcal{E}_{x2}^+ \,,$$

where

$$k_2^+ \equiv n_2^+ \left\{ \cos\theta_2^+ - (E_{z2}^+ / E_{x2}^+) \sin\theta_2^+ \right\} \,. \tag{8A}$$

I have chosen the "positively" traveling wave in medium 2 for illustratory purposes in (6A), (7A), and (8A). Analogous expressions exist for all of the other waves. When the boundary conditions are applied to the magnetic fields we have

$$k_1^+ E_1^+ + k_1^- E_1^- = k_2^+ E_2^+ + k_2^- E_2^- \,, \tag{9A}$$

$$k_2^+ E_2^+ \exp(-j\delta_2^+ d_2) + k_2^- E_2^- \exp(-j\delta_2^- d_2)$$

$$= k_3^+ E_3^+ \exp(-j\delta_3^+ d_2) + k_3^- E_3^- \exp(-j\delta_3^- d_2) \,, \tag{10A}$$

$$k_3^+ E_3^+ \exp(-j\delta_3^+ (d_2+d_3)) + k_3^- E_3^- \exp(-j\delta_3^- (d_2+d_3))$$

$$= k_4^+ E_4^+ \exp(-j\delta_4^+ (d_2+d_3)) \,, \tag{11A}$$

where an x subscript is understood to be affixed to each electric field component as in eqtns. (3A)-(5A).

Before jumping to solutions, a few remarks are in order. Eqtns. (3A), (4A), (5A), (9A), (10A), and (11A) constitute six equations in the six unknowns E^+_{xq}, q = 2,3,4 and E^-_{xq}, q = 1,2,3. Note that the incident wave is assumed to have known characteristics, therefore, E^+_{x1} is not an unknown. The solutions of primary interest in this work are those which describe the reflected wave in medium 1 and the transmitted wave in medium 4. The final solutions will be in the forms E^-_{x1} / E^+_{x1} (tangential reflectance) and E^+_{x4} / E^+_{x1} (tangential transmittance). Knowledge of the x components is sufficient to completely specify all of the electric and magnetic fields for the reflected and transmitted waves. E^+_{y4} / E^+_{x4}, for example, is fixed and H^+_{y4} can be found from an eqtn. similar to (7A). While they do not appear explicitly, the fundamental relations[1],

$$n_1^+ \sin\theta_1^+ = \begin{cases} n_q^+ \sin\theta_q^+ , & q = 2,3,4, \\ n_q^- \sin\theta_q^- , & q = 1,2,3, \end{cases} \tag{12A}$$

have been invoked to obtain the six equations mentioned above. It is of interest to note here that the common rule, "The angle of incidence equals the angle of reflection," does not apply; in general $\sin\theta_q^+ \neq \sin\theta_q^-$, q = 1,2,3, when the ϕ_q are arbitrary.

The solutions for the tangential reflectance and transmittance can, no doubt, be written in a variety of ways, however, for reasons which will be clear later, I prefer the following:

$$\frac{E^+_{x4}}{E^+_{x1}} = \frac{t_{12} T_{24} \exp(j\delta_4^+ d_2)}{\exp(j\delta_2^+ d_2) - r_{21} R_{24} \exp(j\delta_2^- d_2)} , \tag{13A}$$

$$\frac{E^-_{x1}}{E^+_{x1}} = \frac{r_{12}\exp(j\delta_2^+ d_2) + (r_{12}+t_{21}) R_{24} \exp(j\delta_2^- d_2)}{\exp(j\delta_2^+ d_2) - r_{21} R_{24} \exp(j\delta_2^- d_2)} , \tag{14A}$$

where

$$T_{24} = \frac{t_{23}t_{34}\exp(j\delta_4^+ d_3)}{\exp(j\delta_3^+ d_3) - r_{32}r_{34}\exp(j\delta_3^- d_3)}, \tag{15A}$$

$$R_{24} = \frac{r_{23}\exp(j\delta_3^+ d_3) + (r_{23}+t_{32})r_{34}\exp(j\delta_3^- d_3)}{\exp(j\delta_3^+ d_3) - r_{32}r_{34}\exp(j\delta_3^- d_3)}, \tag{16A}$$

The two media tangential reflectances and transmittances are defined by

$$\begin{aligned}
r_{12} &= (k_1^+ - k_2^+)/(k_2^+ - k_1^-) \quad ; \quad t_{12} = (k_1^+ - k_1^-)/(k_2^+ - k_1^-), \\
r_{23} &= (k_2^+ - k_3^+)/(k_3^+ - k_2^-) \quad ; \quad t_{23} = (k_2^+ - k_2^-)/(k_3^+ - k_2^-), \\
r_{34} &= (k_3^+ - k_4^+)/(k_4^+ - k_3^-) \quad ; \quad t_{34} = (k_3^+ - k_3^-)/(k_4^+ - k_3^-), \\
r_{21} &= (k_2^- - k_1^-)/(k_1^- - k_2^+) \quad ; \quad t_{21} = (k_2^- - k_2^+)/(k_1^- - k_2^+), \\
r_{32} &= (k_3^- - k_2^-)/(k_2^- - k_3^+) \quad ; \quad t_{32} = (k_3^- - k_3^+)/(k_2^- - k_3^+).
\end{aligned} \tag{17A}$$

By examining (13A)-(16A), one can perceive a pattern which would allow extension of the solutions to include any number of plates (multilayers) without having to go through the labor involved in solving simultaneous equations. The multilayer problem could also be handled by an adaptation of the matrix techniques devised by Partovi[33].

Eqtns. (5) and (6) in Chapt. I can readily be obtained from (13A)-(16A). The convenient way to achieve this is to consider that media 3 and 4 are identical and then replace all 4 subscripts by 3 subscripts, e.g., $E_{x4}^+ \to E_{x3}^+$. When this is done we find that $r_{34} \to 0$, $t_{34} \to 1$, $R_{24} \to r_{23}$, $T_{24} \to t_{23}$, and (5) and (6) appear from (13A) and (14A), respectively.

The solutions (13A) and (14A) are not yet complete; auxiliary relations must be obtained for the θ_q^+ , n_q^+ , k_q^+ , q = 2,3,4 and θ_q^- , n_q^- , k_q^- , q = 1,2,3. In the work to follow we shall, at times, omit the q subscripts, which identify the medium under consid-

eration, and the $+$, $-$ superscripts, which distinguish the "positively" and "negatively" running waves. Further, it will be convenient to use the following definitions:

$$\epsilon = n^2 \quad , \quad \epsilon_\alpha = n_\alpha^2 \quad , \quad \epsilon_\gamma = n_\gamma^2 \quad , \quad N_1 = (n_1^+ \sin\theta_1^+)^2 \quad ,$$

$$\epsilon_o = \epsilon_\alpha + N_1\left[1 - (\epsilon_\alpha/\epsilon_\gamma)\right] \quad , \quad \epsilon_{qo} = \epsilon_\gamma + N_1\left[1 - (\epsilon_\gamma/\epsilon_\alpha)\right] \quad .$$

ϵ_α and ϵ_γ can be called principal dielectric constants, while ϵ is the dielectric constant of a wave.

The dielectric constant ϵ must satisfy[34]

$$\epsilon\left(\epsilon_\alpha S_\alpha^2 + \epsilon_\gamma S_\gamma^2\right) = \epsilon_\alpha \epsilon_\gamma \quad , \tag{18A}$$

where S_α and S_γ are the α and γ components, respectively, of the unit wave normal $\hat{S}$ and are given by

$$S_\alpha = \sin(\theta - \phi) \quad , \quad S_\gamma = \cos(\theta - \phi) \quad . \tag{19A}$$

Equations (12A) and (18A) constitute two equations in the two unknowns ϵ and θ . Note that θ_1^+ is considered to be known, hence, n_1^+ (or ϵ_1^+) is given directly by (2) in Chapt. I.

Elimination of ϵ from (12A) and (18A) yields

$$\cos[2(\theta - \nu)] = \frac{\epsilon_\alpha \epsilon_\gamma - N_1(\epsilon_\alpha + \epsilon_\gamma)}{\left[\left(\epsilon_o \epsilon_\gamma \cos\phi\right)^2 + \left(\epsilon_{qo} \epsilon_\alpha \sin\phi\right)^2\right]^{1/2}} \quad , \tag{20A}$$

where

$$\tan 2\nu = \frac{2(\epsilon_\alpha - \epsilon_\gamma) N_1}{\epsilon_o \epsilon_\gamma \cot\phi + \epsilon_{qo} \epsilon_\alpha \tan\phi} \quad . \tag{21A}$$

In the qth medium, both θ_q^+ and θ_q^- are found from (20A). The two roots θ_q^+ and θ_q^- are easily identified, however, because $0^\circ \leqq \theta_q^+ \leqq 90^\circ$ and $90^\circ \leqq \theta_q^- \leqq 180^\circ$. The conditions for $180^\circ - \theta_q^- = \theta_q^+$ (angle of incidence equals angle of reflection) can be discerned from (20A) and (21A).

After determining θ , the corresponding n can be found from (12A). Alternatively, a direct analytical expression for n can be used. Elimination of θ from (12A) and (18A) yields

$$\left[\epsilon_\gamma \cos^2\phi + \epsilon_\alpha \sin^2\phi\right] n^2 = \epsilon_0 \epsilon_\gamma^2 \cos^2\phi + \epsilon_{90}\epsilon_\alpha^2 \sin^2\phi$$
$$\pm\left[(\epsilon_\alpha - \epsilon_\gamma)^2 \epsilon_\alpha \epsilon_\gamma N_1 \sin^2 2\phi \cdot \left(\epsilon_\gamma \cos^2\phi + \epsilon_\alpha \sin^2\phi - N_1\right)\right]^{1/2} . \qquad (22A)$$

Unfortunately, the + (-) sign in (22A) does not necessarily go with $n_\beta^+ (n_\beta^-)$, although both n_β^+ and n_β^- are found from (22A). Some information apparently was lost after a squaring operation was performed enroute to (22A). If θ_β^+ and θ_β^- are found first and if $\sin\theta_\beta^+ > \sin\theta_\beta^-$, then (12A) requires that $n_\beta^+ < n_\beta^-$, thus, the negative sign in (22A) corresponds to n_β^+ . This conclusion is reversed if $\sin\theta_\beta^+ < \sin\theta_\beta^-$. The significance of ϵ_0 and ϵ_{90} is apparent from (22A); when $\phi = 0^\circ$ we have $\epsilon = \epsilon_0$ (single solution for both ϵ_β^+ and ϵ_β^-) and when $\phi = 90^\circ$, $\epsilon = \epsilon_{90}$.

The last remaining task is to determine k , which is used in (17A) and defined by (8A). For a wave traveling in an anisotropic dielectric, the electric flux density $\bar{D}$ is related to the electric field E by[34]

$$\bar{D} = \kappa_0 \epsilon \left[\bar{E} - \hat{s}(\hat{s}\cdot\bar{E})\right] , \qquad (23A)$$

where κ_0 is the permittivity of vacuum. By expanding (23A) in the $\alpha\gamma$ reference frame we obtain, for the extraordinary waves,

$$E_\gamma / E_\alpha = \left(\epsilon s_\alpha^2 - \epsilon_\alpha\right) / \epsilon s_\alpha s_\gamma , \qquad (24A)$$

where E_α and E_γ are the α and γ components, respectively, of the vector $\bar{E}$. From a transformation between the xz and $\alpha\gamma$ reference frames we have

$$E_z / E_x = \left\{(E_\gamma/E_\alpha) - \tan\phi\right\} \Big/ \left\{1 + (E_\gamma/E_\alpha)\tan\phi\right\} . \qquad (25A)$$

Now the k factors can be obtained by substituting (24A) and (25A) into the definition for k, which is

$$k \equiv n\{\cos\theta - (E_z/E_x)\sin\theta\} .$$

When this is done, the result can be tidied up by using (18A), yielding

$$\frac{k}{n} = \frac{\epsilon_\gamma \cos^2(\theta-\phi) + \epsilon_\alpha \sin^2(\theta-\phi)}{\epsilon_\gamma \cos(\theta-\phi)\cos\phi - \epsilon_\alpha \sin(\theta-\phi)\sin\phi} . \qquad (26A)$$

The solution of the birefringent plate boundary value problem is now complete in the sense that all formulae necessary for examining the characteristics of the reflected and transmitted waves have been derived. It should be remembered that, up to this point, all of the discussion has concerned waves which have electric fields parallel to the plane of incidence, or extraordinary waves. The solutions (13A) and (14A) for the tangential transmittance and reflectance can be applied to ordinary waves, however, as shown in the following paragraphs.

If the incident wave is of the ordinary kind, it has an electric field which is perpendicular to the plane of incidence and an x component of magnetic field given by (MKS units)

$$j\omega\mu_0 \mathcal{H}_x = \partial \mathcal{E}_y/\partial z ,$$

or

$$(-\omega\mu_0\lambda/2\pi)\,\mathcal{H}_x = n\cos\theta\,\mathcal{E}_y .$$

For the ordinary waves, the y components of electric field and x components of magnetic field must be continuous across the boundaries, hence, if we define k by

$$k \equiv n\cos\theta , \qquad (27A)$$

then the six equations (3A), (4A), (5A), (9A), (10A) and (11A), where a y subscript is understood, e.g., $E_1^+ \rightarrow E_{y1}^+$, are valid for the ordinary waves.

Various simplifications arise in the case of ordinary waves. In the qth medium both "positively" and "negatively" traveling waves have the same index of refraction or

$$n_q^+ = n_q^- = n_{\beta q} = n_{yq} \,. \tag{28A}$$

Also, "the angle of incidence equals the angle of reflection," so that

$$\sin\theta_q^+ = \sin\theta_q^- \,, \quad \cos\theta_q^+ = -\cos\theta_q^- \,,$$

$$\underline{\text{or}} \quad \theta_q^+ = \pi - \theta_q^- \equiv \theta_q \,. \tag{29A}$$

The implications of (28A) and (29A) are that

$$k_q^+ = -k_q^- \equiv k_q = n_{yq}\cos\theta_q \,,$$

$$\delta_q^+ = -\delta_q^- \equiv \delta_q = 2\pi n_{yq}\cos\theta_q/\lambda \,. \tag{30A}$$

Using (30A), the solutions for the tangential transmittance and reflectance simplify to

$$\frac{E_{y4}^+}{E_{y1}^+} = \frac{t_{12}T_{24}\exp(j\delta_4 d_2 - j\delta_2 d_2)}{1 + r_{12}R_{24}\exp(-j2\delta_2 d_2)} \,, \tag{31A}$$

$$\frac{E_{y1}^-}{E_{y1}^+} = \frac{r_{12} + R_{24}\exp(-j\delta_2 2d_2)}{1 + r_{12}R_{24}\exp(-j2\delta_2 d_2)} \,, \tag{32A}$$

$$T_{24} = \frac{t_{23}t_{34}\exp(j\delta_4 d_3 - j\delta_3 d_3)}{1 + r_{23}r_{34}\exp(-j2\delta_3 d_3)} \,, \tag{33A}$$

$$R_{24} = \frac{r_{23} + r_{34}\exp(-j2\delta_3 d_3)}{1 + r_{23}r_{34}\exp(-j2\delta_3 d_3)} \,. \tag{34A}$$

The two media transmittances and reflectances are given by

$$r_{12} = (k_1 - k_2)/(k_1 + k_2) \;;\; t_{12} = 2k_1/(k_1 + k_2) \,,$$

$$r_{23} = (k_2 - k_3)/(k_2 + k_3) \;;\; t_{23} = 2k_2/(k_2 + k_3) \,, \tag{35A}$$

$$r_{34} = (k_3 - k_4)/(k_3 + k_4) \;;\; t_{34} = 2k_3/(k_3 + k_4) \,.$$

Eqtns (31A)-(35A) are identical to those obtained for isotropic media and can be found in a number of texts on optics and thin films; they have been included in this appendix mainly for convenient reference.

APPENDIX B

Exact Theory of Retardation Plates

This appendix is a self contained work and, in fact, has been published in the Journal of the Optical Society of America, vol. 54, p. 115, 1964. The differences between this appendix and the published paper of the same title are only minor. Permission to reproduce the published work has been granted by the Journal of the Optical Society of America.

Exact Theory of Retardation Plates*

D. A. Holmes
Department of Electrical Engineering
Carnegie Institute of Technology
Pittsburgh 13, Pennsylvania

Abstract

The conventional solutions to the problem of normally incident light transmission through homogeneous, birefringent, non-optically active, non-absorbing, crystalline plate are not exact. When treated as a boundary value problem in electromagnetic field theory, exact expressions are obtained for the retardation or phase difference and the electrical field amplitude ratio. The two solutions differ in some interesting ways that become of substantial importance in the examination of laser light. The nature of the exact solutions is examined in detail and numerical comparisons with the conventional solutions are given for the cases of calcite and quartz, neglecting the optical activity of crystalline quartz. For quartz it is shown that one can obtain a quarter-wave plate by using any one of a number of different crystal thicknesses. The application of wave plates in the investigation of elliptically polarized light is briefly discussed. For small angles of incidence, the effects to be expected for light which is obliquely incident on the plate surface are investigated.

*This work was supported by the National Aeronautics and Space Administration under contract NAS8-5269.

INTRODUCTION

It is not generally acknowledged in the scientific literature that the commonly used expressions which describe the polarization state of an electromagnetic wave emerging from an anisotropic dielectric plate are approximations which have been derived by neglecting the effect of multiple internal reflections between the plate surfaces. To review briefly the conventional theory let us consider Fig. B-1. An anisotropic dielectric plate of thickness d has been oriented with respect to an xyz coordinate system such that the principal axes of the plate are aligned with the coordinate axes, with the xy plane coinciding with one of the crystal surfaces. The index of refraction in the x direction is denoted by n_x, in the y direction by n_y and in the z direction by n_z. It is considered that a monochromatic plane wave of angular frequency ω and traveling in the positive z direction is normally incident on the left crystal surface. The components of the incident electric field are then given by

$$\mathcal{E}_{ix} = E_{ix} e^{j\tau} , \tag{1a}$$

$$\mathcal{E}_{iy} = E_{iy} e^{j\tau} , \tag{1b}$$

where $\tau \equiv \omega t - \beta_o z$, t is time, $\beta_o = 2\pi/\lambda_o$, and λ_o is the vacuum wavelength. At this time E_{ix} and E_{iy} are considered to be real numbers. The traditional geometrical optics solutions[1] for the components of the transmitted electric field are then given by

$$\mathcal{E}_{ox} = E_{ox} e^{j[\tau - (\beta_x - \beta_o)d]} , \tag{2a}$$

$$\mathcal{E}_{oy} = E_{oy} e^{j[\tau - (\beta_y - \beta_o)d]} , \tag{2b}$$

[1] For an equivalent analysis see, for example, M. Born and E. Wolf, Principles of Optics (Pergamon Press, Inc., New York, 1959), pp. 688 et seq.

where $\beta_x = 2\pi n_x/\lambda_o$ and $\beta_y = 2\pi n_y/\lambda_o$. The phase difference between the x and y components is given by

$$\Delta_a = (\beta_y - \beta_x)d. \tag{3}$$

The subscript a may be taken to imply that Δ_a has been calculated by approximate theory. It is to be noted that, for the approximate theory, the amplitude factors of the x and y components of the transmitted wave are respectively set equal to the x and y components of the incident wave, i.e., $E_{ox} = E_{ix}$ and $E_{oy} = E_{iy}$, and that the phase difference Δ_a depends linearly on the crystal thickness.

Stephanov and Khapalyuk[2] have shown that the approximations of geometrical optics allow only a qualitative description of the transmission characteristics of an isotropic plate with a negative absorption coefficient. We shall show that geometrical optics gives only an approximate account of transmission through a lossless, birefringent, plane-parallel layer. F. Gabler and P. Sokob[3] have considered the effect of reflections on the transmission characteristics of doubly refracting plates, however, they did not compare their solutions with the conventional solutions and did not give numerical results for actual crystals. In addition, the reflexionskoeffizient r_i used by them did not take into account the anisotropy of the crystal. Therefore, it is the purpose of this work to give an account of the nature of the rigorous wave optics solutions, using calcite and quartz as numerical examples.

ISOTROPIC PLATE

It is instructional to consider some aspects of plane wave propagation through an isotropic, i.e., $n_x = n_y = n_z = n$, plate. If the incident and transmitted electric fields are respectively written as

---

[2] B. I. Stepanov and A. P. Khapalyuk, Optics and Spectroscopy, 13, 404 (1962).

[3] F. Gabler and P. Sokob, Z Physik, 116, 47 (1940).

$$\mathcal{E}_i = E_i e^{j\tau} , \quad \mathcal{E}_o = E_o e^{j[\tau - \theta + \beta_o d]} ,$$

where E_i and E_o are considered to be real numbers, then, by an analysis similar to that of Fry[4] or that of Jacobs, et al.[5], one can find that

$$\frac{E_o}{E_i} = \left[\cos^2 \beta d + K^2 \sin^2 \beta d \right]^{-1/2} , \tag{4}$$

$$\theta = \arctan\left[K \tan \beta d \right] , \tag{5}$$

$$K = \frac{n^2+1}{2n} , \quad \beta = 2\pi n / \lambda_o . \tag{6}$$

Alternatively, one can derive the above relationships for the amplitude transmission coefficient E_o / E_i and the phase shift $\beta_o d - \theta$ from transmission line theory, using the concept of wave impedance.[6] The transmission line approach has been applied successfully in microwave[7] and optical[8] problems involving geometries similar to that used in the present work. One can determine the physical implications of (4) and (5) in a relatively straight-forward manner. For example, when the plate thickness is equal to an integer multiple of half-wavelengths, $d = m\lambda_o / 2n$, $m = 0, 1, 2, \ldots$, the input wave impedance seen by the incident wave at the $z = 0$ plane is matched to the intrinsic wave impedance of air and the plate acts as a transparent half-wave

[4] Thornton C. Fry, J. Opt. Soc. Am. and Rev. Sci. Instr. 16, 1 (1928).

[5] H. Jacobs, D. A. Holmes, L. Hatkin, and F. A. Brand, J. Appl. Phys. 34, 2617 (1963).

[6] S. Ramo and J. R. Whinnery, Fields and Waves in Modern Radio (John Wiley and Sons, Inc., New York, 1960), pp. 290 et seq.

[7] H. Jacobs, F. A. Brand, J. D. Meindl, S. Weitz, R. Benjamin, and D. A. Holmes, Proc. IEEE 51, 581 (1963).

[8] H. Jacobs, D. A. Holmes, L. Hatkin and F. A. Brand, "Transmission Line Formulation for Optical Maser Amplification" (USAELRDL, Fort Monmouth, N. J., Technical Report 2402, Nov. 1963). Also J. Opt. Soc. Am. 54, 1416 (1964).

dielectric window. Equivalently, one can say that the reflected wave is zero because each contribution from an internal reflection has such phase characteristics that, when the infinite summation of all contributions is taken, the net reflection is zero. For the half-wave window, $\theta = m\pi$. For any other value of thickness, the input impedance seen at $z = 0$ is different from the intrinsic impedance of the input medium (air). Because of this impedance mismatch, the plate acts as a semi-transparent window which reflects a fraction of the incident radiation. In particular, the plate acts as an anti-window when the thickness is an odd multiple of quarter-wavelengths, or when $d = (2m+1)\lambda_0/4n$, $m = 0, 1, 2, 3, \ldots$. For this case a minimum in the amplitude transmission coefficient occurs and $\theta = (2m+1)\pi/2$. The minimum in transmission coefficient is given by $E_0/E_i = [2n/(n^2+1)]^{1/2}$.

If we define an error angle e by

$$e \equiv \arctan[K \tan\beta d] - \beta d \,, \tag{7}$$

then it is seen that e is the difference between the approximate and exact solutions for the phase constant of the transmitted wave. The quantity e is periodic in d , experiencing finite maxima and minima. By maximizing (7) with respect to d it is found that

$$e_{max} = \arctan\left\{\frac{K-1}{2\sqrt{K}}\right\} \,. \tag{8}$$

In Fig. B-2 is shown a curve of e_{max} versus index of refraction n. It is seen that the maximum error introduced by using the approximate solution for the phase constant of the transmitted wave increases with the refractive index of the isotropic plate.

For later purposes, let us now derive an approximate form for e which is valid for $1 \leq n < 2$. Rewriting (7) as

$$e = \tan^{-1}[\tan\beta d + (K-1)\tan\beta d] - \tan^{-1}[\tan\beta d] \,,$$

we recognize that, for $(K-1) \ll 1$, we may approximate e by

$$e \cong (K-1)\tan\beta d/(1+\tan^2\beta d) = \frac{K-1}{2}\sin 2\beta d . \qquad (9)$$

For $n=2$ we have $e_{max} \cong (K-1)/2 \cong 0.125$ radian $\cong 7.2^\circ$, which is in fair agreement with Fig. 2. With the aid of (9), θ can be approximated by

$$\theta \cong \beta d + \frac{K-1}{2}\sin 2\beta d . \qquad (10)$$

We note that, in this approximation, when the plate is an eighth-wave semi-transparent window we have the largest discrepancy between the exact and approximate solutions for θ . When $d = (2m+1)\lambda_0/8n$, $m = 0,1,2,\ldots$, we see that $\theta \cong (2m+1)\pi/4 \pm (K-1)/2$, where the (+) sign goes with even values of m and the (-) sign goes with odd values of m . We shall hereafter speak of an even (odd) $\lambda/8$ plate when the plus (minus) sign is applicable.

ANISOTROPIC PLATE

For a given value of thickness, the anisotropic plate will exhibit polarization dependent window properties. Suppose, for example, that the plate thickness is such that the plate simultaneously acts approximately as an odd $\lambda/8$ plate for the x component of incident electric field and an even $\lambda/8$ plate for the y component. Then the phase difference between the emergent x and y components will differ from the value predicted by approximate theory by an amount equal to the sum of the absolute values of the discrepancies between exact and approximate theory for the respective x and y phase shifts. If the plate simultaneously acts approximately as a transparent window for one component and as an anti-window for the other component, then it is clear that the ratio of the y amplitude to the x amplitude for the transmitted wave will differ from that ratio for the incident wave. Both of these features will now be examined from a quantitative standpoint.

The results of the previous section can be easily generalized to the case of a birefringent plate for the geometry of Fig. B-1. The transmitted electric field components for the anisotropic plate are

$$\mathcal{E}_{ox} = \frac{E_{ix}\, e^{j[\tau - \theta_x + \beta_o d]}}{\left[\cos^2\beta_x d + K_x^2 \sin^2\beta_x d\right]^{1/2}}, \tag{11}$$

$$\mathcal{E}_{oy} = \frac{E_{iy}\, e^{j[\tau - \theta_y + \beta_o d]}}{\left[\cos^2\beta_y d + K_y^2 \sin^2\beta_y d\right]^{1/2}}, \tag{12}$$

where

$$K_k = \frac{n_k^2 + 1}{2n_k}, \quad \theta_k = \tan^{-1}\left[K_k \tan\beta_k d\right]; \quad k = x, y.$$

The phase difference between the x and y components is now given by

$$\Delta_e = \theta_y - \theta_x = \arctan\left[\frac{K_y \tan\beta_y d - K_x \tan\beta_x d}{1 + K_x K_y \tan\beta_x d \tan\beta_y d}\right], \tag{13}$$

while the amplitude ratio is

$$\frac{E_{oy}}{E_{ox}} = \frac{E_{iy}}{E_{ix}} \cdot F \equiv \frac{E_{iy}}{E_{ix}} \cdot \left[\frac{\cos^2\beta_x d + K_x^2 \sin^2\beta_x d}{\cos^2\beta_y d + K_y^2 \sin^2\beta_y d}\right]^{1/2}. \tag{14}$$

To gain an understanding of the way Δ_e compares with Δ_a, in Fig. B-3 we have plotted the difference $\Delta_e - \Delta_a$ versus d/λ_o for a calcite crystal using $n_y = 1.64869$ and $n_x = 1.48216$. The values of n_y and n_x correspond to a wavelength of 8010 Angstroms.[9] From inspection of Fig. B-3, it would appear that the error angle $\Delta_e - \Delta_a$ has an amplitude modulated sinusoidal variation with crystal thickness. Indeed, this is approximately the case, as will

now be shown. If we write

$$\Delta_e - \Delta_a = \left\{\tan^{-1}\left[\tan\beta_y d + (K_y - 1)\tan\beta_y d\right] - \tan^{-1}\left[\tan\beta_y d\right]\right\}$$

$$- \left\{\tan^{-1}\left[\tan\beta_x d + (K_x - 1)\tan\beta_x d\right] - \tan^{-1}\left[\tan\beta_x d\right]\right\},$$

then by using the result (9), one can obtain

$$\Delta_e - \Delta_a \cong \frac{K_y - 1}{2}\sin 2\beta_y d - \frac{K_x - 1}{2}\sin 2\beta_x d,$$

or

$$\Delta_e - \Delta_a \cong \frac{K_y + K_x - 2}{2} \cdot \cos\left[(\beta_y + \beta_x)d\right] \cdot \sin\Delta_a . \tag{15}$$

For calcite we have that

$$\frac{K_y + K_x - 2}{2} = \frac{(n_x + n_y)(1 + n_x n_y) - 4n_x n_y}{4n_x n_y} \cong 0.105 \text{ rad} \cong 6^\circ .$$

The maximum excursion of $\Delta_e - \Delta_a$ from zero is about 5.5° as determined from Fig. B-3, which is in good agreement with the approximation (15).

Next we consider the shift in the amplitude ratio which is given by F in (14). For calcite, F is plotted versus d/λ_o in Fig. B-4. It is noted that F can differ from unity by as much as 10% for the exact theory, while F is considered to be unity for all crystal thicknesses in the approximate theory. A simple approximation for F can be obtained as follows.

$$F = \left[\frac{1 + (K_x^2 - 1)\sin^2\beta_x d}{1 + (K_y^2 - 1)\sin^2\beta_y d}\right]^{1/2},$$

$$F \cong 1 - \frac{K_y^2 + K_x^2 - 2}{4} \cdot \sin[(\beta_y + \beta_x)d] \cdot \sin \Delta_a$$

$$+ \frac{K_x^2 - K_y^2}{4} \cdot \left\{ 1 - \cos[(\beta_y + \beta_x)d] \cdot \cos \Delta_a \right\} . \qquad (16)$$

Since $K_y > K_x$ one can see that (16) would tend to oscillate about a value less than unity. This behavior can be detected by examining Fig. B-4.

Regarding the general oscillatory nature of $\Delta_e - \Delta_a$ and F for calcite one observes that the period of the slowly varying factors is given by $(n_y - n_x)^{-1} \cong 6$ and that the period of the rapidly varying factors is $(n_y + n_x)^{-1} \cong 0.32$.

In Fig. B-5, Δ_e and Δ_a are both plotted for values of phase shift in the range 72°-106°. For certain values of phase shift, e.g. 82°, it can be noted that three different crystal thicknesses would yield identical values of phase shift, according to exact theory.

Turning now to quartz[9], for $\lambda_o = 8325\text{Å}$, we have $n_x = 1.53773$ and $n_y = 1.54661$. Now the rapidly varying factor in (15) has a period $(n_x + n_y)^{-1} \cong 0.325$, which is nearly the same period as that for calcite. The period of the slowly varying factor for quartz is $(n_x - n_y)^{-1} \cong -113$, which is over an order of magnitude larger than the corresponding number for calcite. In Fig. B-6 the exact phase shift Δ_e is compared with the approximate phase shift Δ_a for thickness values in the range $26.4 \leqq d/\lambda_o \leqq 30.0$. It is seen that Δ_e oscillates about Δ_a in a regular fashion with the maximum difference between Δ_e and Δ_a being somewhat greater than 5°. Further it is seen that 21 different crystal thicknesses can be used to obtain a phase shift of 90°, according to the exact theory. Fig. B-7 shows the shift in amplitude ratio F as a function of d/λ_o. It should be observed

---

[9] For the purposes of numerical computations we have chosen optical constants for calcite and quartz which correspond to near infrared wavelengths because we are using GaAs injection laser sources (8400°A) in some of our work. The general concepts advanced in this work, however, are valid for other wavelengths. The numerical values for the optical constants were taken from: Dwight E. Gray, Coordinating Editor, American Institute of Physics Handbook (McGraw-Hill Book Company, Inc., New York, 2nd Edition, 1963), Calcite, p. 6-18; Crystal Quartz, p. 6-24; Rutile, p. 6-33.

that, for each value of d/λ_o which gives $\Delta_e = 90^o$, the corresponding value of F is different. In closing the above discussion of propagation through quartz, it must be pointed out that crystalline quartz is an optically active substance.[10] Rather than include the gyration vector or the optical activity vector in the Maxwell equation analysis we have chosen to neglect the effects of optical activity in quartz.

To demonstrate in a more vivid fashion the limitations of the geometrical optics approximation let us consider, as a final numerical example, transmission through a rutile crystal at 5770 Å. At this wavelength the optical constants for rutile[9] are $n_y = 2.921$ and $n_x = 2.623$. For $d/\lambda_o = 0.99$ the computer calculations reveal that $\Delta_e = 80.9^o$ and $\Delta_a = 106.2^o$. The discrepancy between the wave optics results and the geometrical optics result is $\Delta_e - \Delta_a = -25.3^o$. For $d/\lambda_o = 0.86$ the electric field amplitude ratio shift is $F = 1.48$. For $d/\lambda_o = 0.77$ we have that $F = 0.62$.

ANALYSIS OF ELLIPTICALLY POLARIZED LIGHT

We now assume that the incident electric field is elliptically polarized such that $E_{iy}/E_{ix} = e^{j\delta}\tan\gamma$. Further we assume that the azimuth of the major axis of the elliptic vibration with respect to the positive x axis is given by ψ and that the ratio of the minor axis of the ellipse to the major axis of the ellipse is given by $\tan\phi$. These quantities are then related by[11]

$$\sin 2\phi = \pm \sin 2\gamma \sin\delta, \tag{17a}$$

$$\tan 2\psi = \tan 2\gamma \cos\delta. \tag{17b}$$

---

[10] G. N. Ramachandran and S. Ramaseshan in S. Flugge, Editor, Handbuch der Physik (Springer-Verlag, Berlin, Band XXV/1 1961), pp. 76 et seq.

[11] Reference 1, pp. 24 et seq.

In the approximate theory, it is generally conceded that ψ and $\tan\phi$ are measurable by a compensator-analyzer combination and, hence, δ and $\tan\gamma$ can then be calculated. If we consider now that compensation is achieved by an exact quarter-wave plate, i.e., $\Delta_e = 90^\circ$ at the wavelength used, then the compensator setting will yield ψ and (17b) can be used. According to the exact theory, the analyzer setting will determine $\tan\nu$ where $\tan\nu = F\tan\phi$, hence (17a), in terms of the measured quantity $\tan\nu$, becomes

$$2F\tan\nu/(F^2+\tan^2\nu) = \pm \sin 2\gamma \sin\delta . \qquad (18)$$

If the compensator is not an exact quarter-wave plate, i.e., $\Delta_e \neq 90^\circ$, then the analysis of Hall[12], or Bergman[13], can be used, however, (18) is now modified to

$$2F\tan\nu \sin\Delta_e/(F^2+\tan^2\nu) = \pm \sin 2\gamma \sin\delta . \qquad (19)$$

Equation (19) has been obtained by making the exact theory correction to the final equation in Hall's paper.

OBLIQUE INCIDENCE

In the previous sections it was assumed that the incident ratiation was a normally incident, ideal plane wave. In an experimental situation when one uses a beam of light it is of interest to be able to assess the

---

[12] A. C. Hall, J. Opt. Soc. Am. 53, 801(1963). We have observed a typographical error in equation (12) of Hall's work. The corrected forms, in Hall's notation are

$$\cos 2\Psi = [+\cos\delta \sin\delta \pm \cot 2\gamma (K\tan\Delta - \sin^2\delta)^{1/2}]/K ,$$

$$\sin 2\Psi = [-\sin\delta \cot 2\gamma \pm \cos\delta (K\tan\Delta - \sin^2\delta)^{1/2}]/K .$$

The geometry used by Hall is equivalent to that used in the present work.

[13] D. Bergman, J. Opt. Soc. Am. 52, 1080 (1962).

importance of a non-zero beam divergence and the effects introduced when the wave-plate surfaces are not perfectly perpendicular to the beam axis. We shall approach the problems of imperfect collimation and wave-plate misalignment by investigating the properties of a plane wave which is transmitted through the wave-plate at oblique incidence. This is tantamount to assuming that a diverging beam of light can be approximated by a summation of plane waves which are traveling in slightly different directions. To simplify the analysis we shall consider that the beam divergence is confined to the xz plane. An incident plane wave traveling in the xz plane at an angle i with respect to the z axis has components given by

$$\mathcal{E}_{ix} = E_{ix}\, e^{j\tau} \quad , \quad \mathcal{E}_{iy} = E_{iy}\, e^{j\tau} \, ,$$

where $\tau = \omega t - \beta_0 (x \sin i + z \cos i)$. The incident field also has a z component which can be related to E_{ix} .

Representing the transmitted components by

$$\mathcal{E}_{ox} = E_{ox}\, e^{j[\tau - \theta_x + \beta_0 d \cos i]} \, ,$$

$$\mathcal{E}_{oy} = E_{oy}\, e^{j[\tau - \theta_y + \beta_0 d \cos i]} \, ,$$

it is then found that the expressions for Δ_a , Δ_e , and F are still given by (3), (13), and (14), respectively, when the following substitutions are made:

$$\beta_y = (2\pi/\lambda_0)(n_y^2 - \sin^2 i)^{1/2} \, , \tag{20}$$

$$\beta_x = (2\pi n_x/\lambda_0 n_z)(n_z^2 - \sin^2 i)^{1/2} \, ,$$

$$K_y = \frac{1}{2}\left\{ \frac{\cos i}{(n_y^2 - \sin^2 i)^{1/2}} + \frac{(n_y^2 - \sin^2 i)^{1/2}}{\cos i} \right\} , \tag{21}$$

$$K_z = \frac{1}{2}\left\{ \frac{n_x n_z \cos i}{(n_z^2 - \sin^2 i)^{1/2}} + \frac{(n_z^2 - \sin^2 i)^{1/2}}{n_x n_z \cos i} \right\} . \tag{22}$$

(23)

In Fig. B-8 we show Δ_e and Δ_a as functions of crystal thickness for two waves, one which is normally incident and the other incident in the xz plane at 5°. It is observed that, if both beams are simultaneously present, the exact theory predicts that it would be virtually impossible to perfectly compensate both beams for a single crystal thickness. This feature is not nearly so apparent in the approximate theory. For example, consider a crystal of thickness $d/\lambda_0 = 28.0$. At this thickness value, Δ_e for the 5° wave is 4° greater than Δ_e for the 0° wave, while Δ_a for the 5° wave is only 0.15° greater than Δ_a for the 0° wave.

In Fig. B-9, Δ_e and Δ_a are shown as functions of the angle of incidence i for some selected crystal lengths. For small angles of incidence such that $\sin^2 i \cong i^2$, Δ_a has a small component proportional to i^2. The slow variation of Δ_a with i is apparent from Fig. B-9. The exact theory, however, indicates that Δ_e has a relatively strong dependence on the angle of incidence for $i > 1^\circ$.

SUMMARY AND CONCLUSIONS

By using the Maxwell equations with appropriate boundary conditions, exact solutions for the phase difference and amplitude shift have been obtained for the case of normal incidence propagation through an anisotropic dielectric plate. Approximations were developed which enable one to visualize more clearly the functional nature of the exact solutions. The exact solutions were numerically compared with the conventional solutions, using for examples the optical constants of calcite and quartz. It was found that the discrepancy between the exact solutions and the conventional solutions could be as much as 5° for the phase difference and as much as 10% for the transmitted amplitude ratio. The limitations of geometrical optics become more apparent when considering propagation through a crystal with relatively large principal refractive indices. In the case of rutile at 5770 Å, the exact and approximate values of phase difference can differ by as much as 25°. For crystals characterized by very small differences in the principal indices of refraction, such as quartz, it was determined that many

different crystal thicknesses can yield identical values of phase difference. When retardation plates are used in the investigation of elliptically polarized light wherein lasers or other highly monochromatic, well collimated sources are used, it is felt that use of the exact relations will be necessitated when accurate results are desired. The exact theory predicts that it is more difficult to compensate all of a diverging beam of light with a waveplate than does the approximate theory. It should be observed that, although the thickness values used in the numerical examples might correspond to physically unrealizable dimensions, the general conclusions reached in this work are still valid because of the periodicity of the exact solutions with crystal thickness.

The writer is pleased to note that, immediately prior to the submission of this work, Weinberger and Harris[14] reported measurements which support some of the theoretically based conclusions reached herein. Portions of their interpretation of the exact equations are in error, however. For example, they stated that, for a perfect quarter-wave plate, the following equations will be satisfied

$$n_x d - (2m\lambda_0/4) = 0, \tag{24a}$$

$$n_y d - (2m\lambda_0/4) = \lambda_0/4, \tag{24b}$$

where m is an integer. If we consider that n_x, n_y, and λ_0 are known, then (24a) and (24b) can simply be treated as two simultaneous equations in the two unknowns, m and d, with solutions

$$m = n_x / [2(n_y - n_x)], \tag{25a}$$

$$d = \lambda_0 / [4(n_y - n_x)]. \tag{25b}$$

From (25a) it is clear that, only under fortuitous circumstances, would m be exactly an integer for a real crystal.

From equation (13) in this work it is determined that 90° phase

---

[14] H. Weinberger and J. Harris, J. Opt. Soc. Am. 54, 552 (1964).

difference will occur for values of d satisfying

$$1 + K_x K_y \tan\beta_x d \tan\beta_y d = 0 . \qquad (26)$$

ACKNOWLEDGEMENTS

The author wishes to acknowledge useful discussions, with support and encouragement, from Drs. D. L. Feucht and R. L. Longini, Carnegie Institute of Technology. The author is indebted to Dr. H. Jacobs, U. S. Army Electronics Research and Development Laboratory, Fort Monmouth, New Jersey, for many helpful and interesting discussions regarding the transmission and reflection properties of isotropic plates. Messrs. J. Valerio and A. Smith, Carnegie Institute of Technology, assisted in the preparation of this work.

Figure Captions for Appendix B

Fig. B-1 Geometry used in this work. Theanisotropic plate is assumed to have its principal dielectric axes aligned with the cartesian coordinate system shown. The x direction is that for a right hand coordinate system.

Fig. B-2 The maximum error angle e_{max}, for an isotropic slab, versus index of refraction n.

Fig. B-3 The quantity $(\Delta_e - \Delta_a)$ versus d/λ_o for calcite at a wavelength of 8010 Angstroms.

Fig. B-4 The shift in the amplitude ratio $F \equiv (E_{oy}/E_{ox})/(E_{iy}/E_{ix})$ versus d/λ_o for calcite at 8010 Angstroms wavelength.

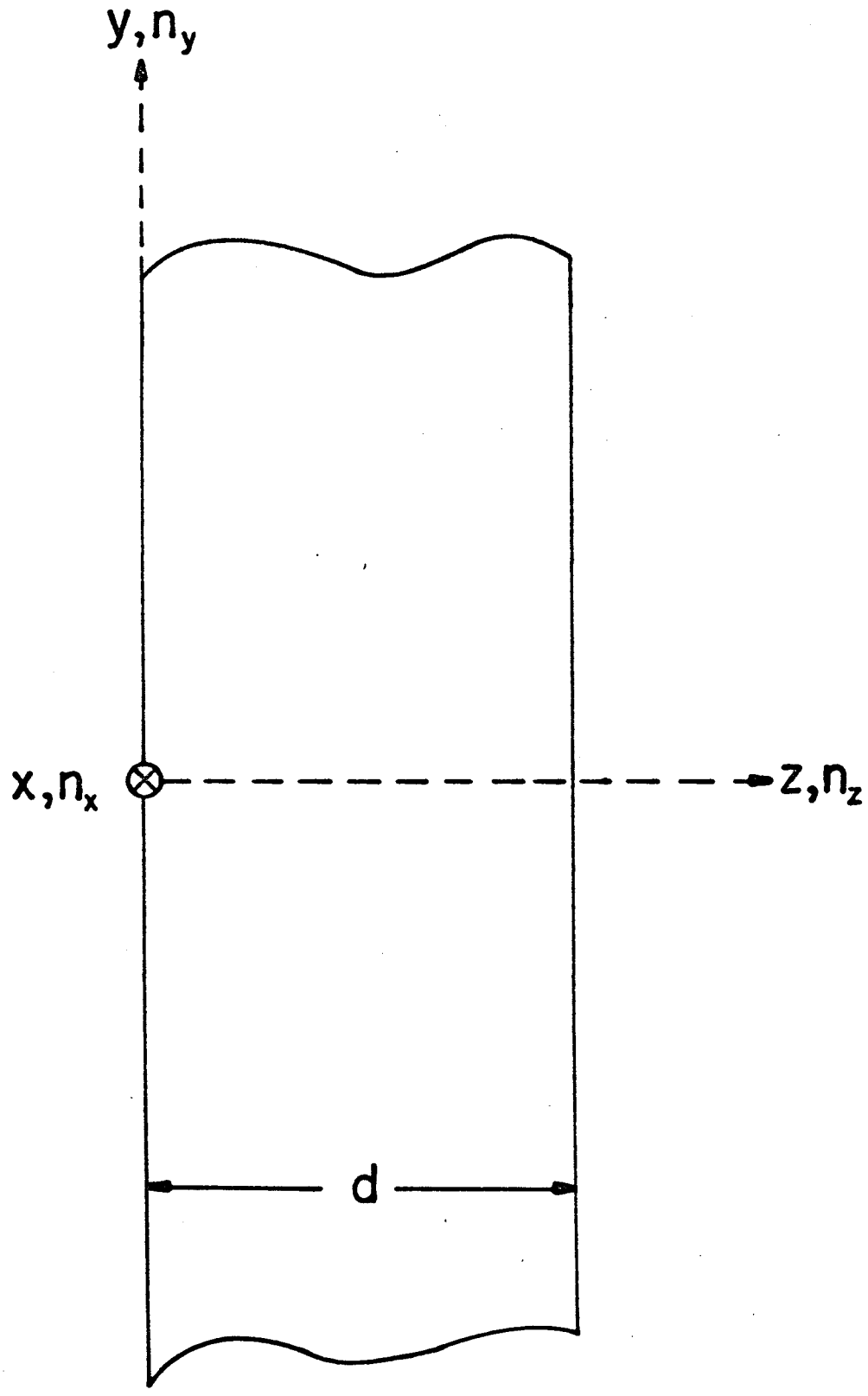
Fig. B-5 Comparison between Δ_e and Δ_a for calcite in the vicinity of 90° phase difference.

Fig. B-6 Comparison between Δ_e and Δ_a for quartz in vicinity of 90° phase difference. The optical constants used are for 8325 Angstroms.

Fig. B-7 The shift in amplitude ratio $F \equiv (E_{oy}/E_{ox})/(E_{iy}/E_{ix})$ versus d/λ_o for quartz at 8325 Angstroms wavelength.

Fig. B-8 Exact (Δ_e) and approximate (Δ_a) values of phase difference for quartz as a function of normalized crystal thickness d/λ_o with angle of incidence i as a parameter. $n_y = 1.54661$. $n_x = n_z = 1.53773$.

Fig. B-9 Δ_e and Δ_a versus angle of incidence for quartz. The numbers arranged in a column on the right correspond to values of d/λ_o. The slowly rising curves represent the approximate phase difference Δ_a. For $\lambda_o = 8325$ Angstroms, $n_x = n_z = 1.53773$, $n_y = 1.54661$.



ANISOTROPIC DIELECTRIC PLATE

Fig. B-1

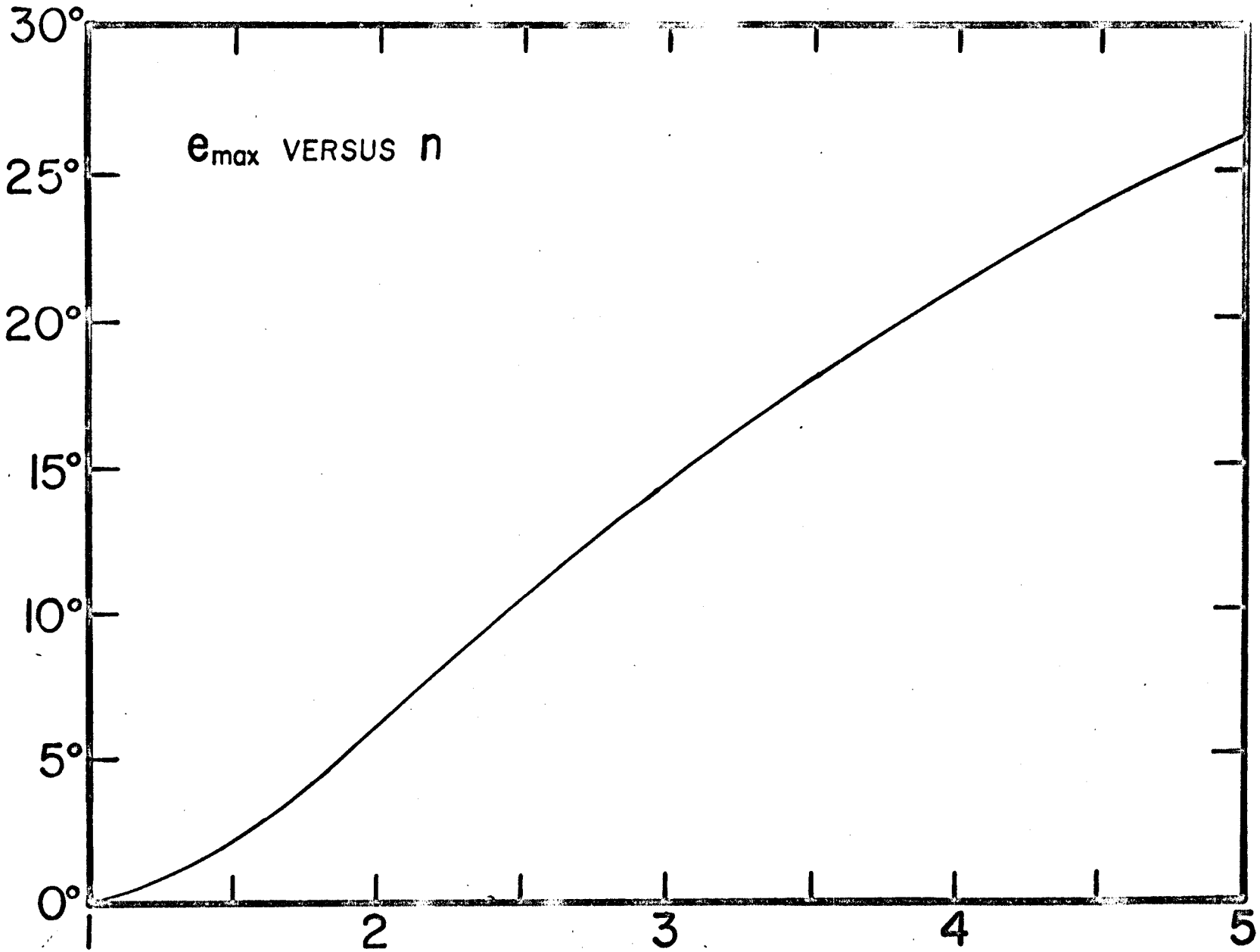


Fig. B-2

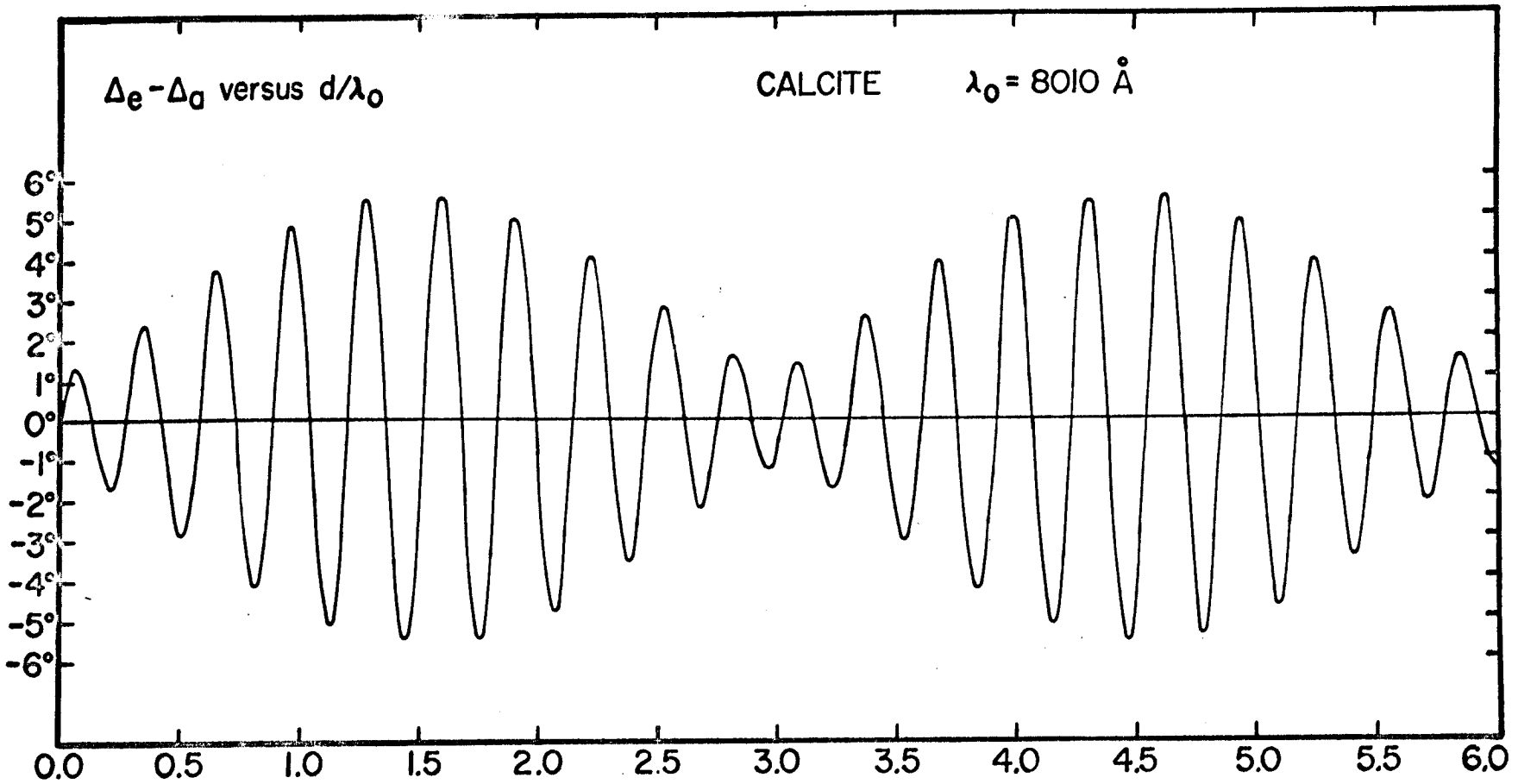

$\Delta_e - \Delta_o$ versus d/λ_o
CALCITE
λ_o = 8010 Å
6°
5°
4°
3°
2°
1°
0°
-1°
-2°
-3°
-4°
-5°
-6°
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0

Fig. B-3

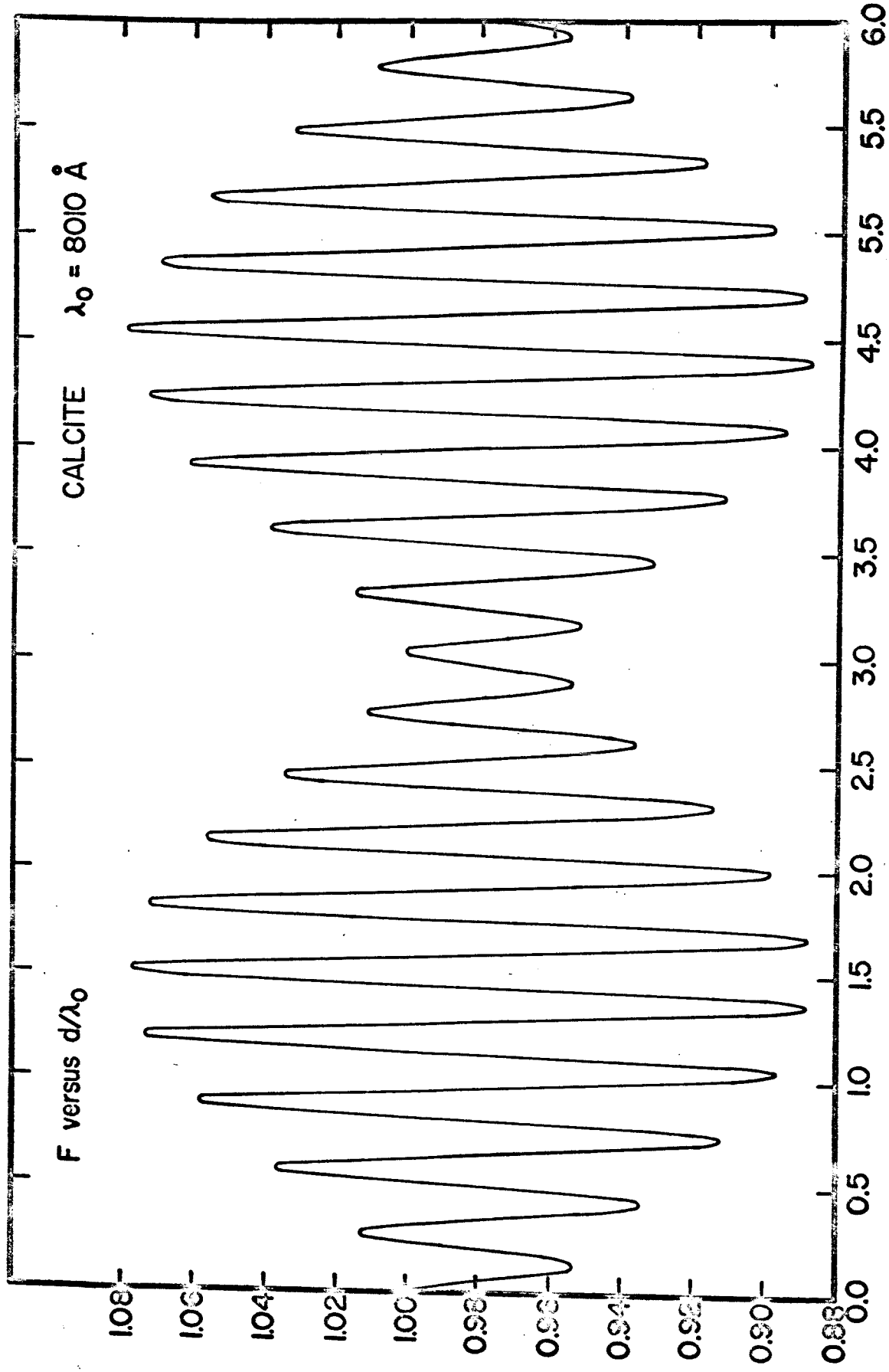

F versus d/λo
CALCITE λo = 8010 Å
1.08
1.06
1.04
1.02
1.00
0.98
0.96
0.94
0.92
0.90
0.88
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0

Fig. B-4

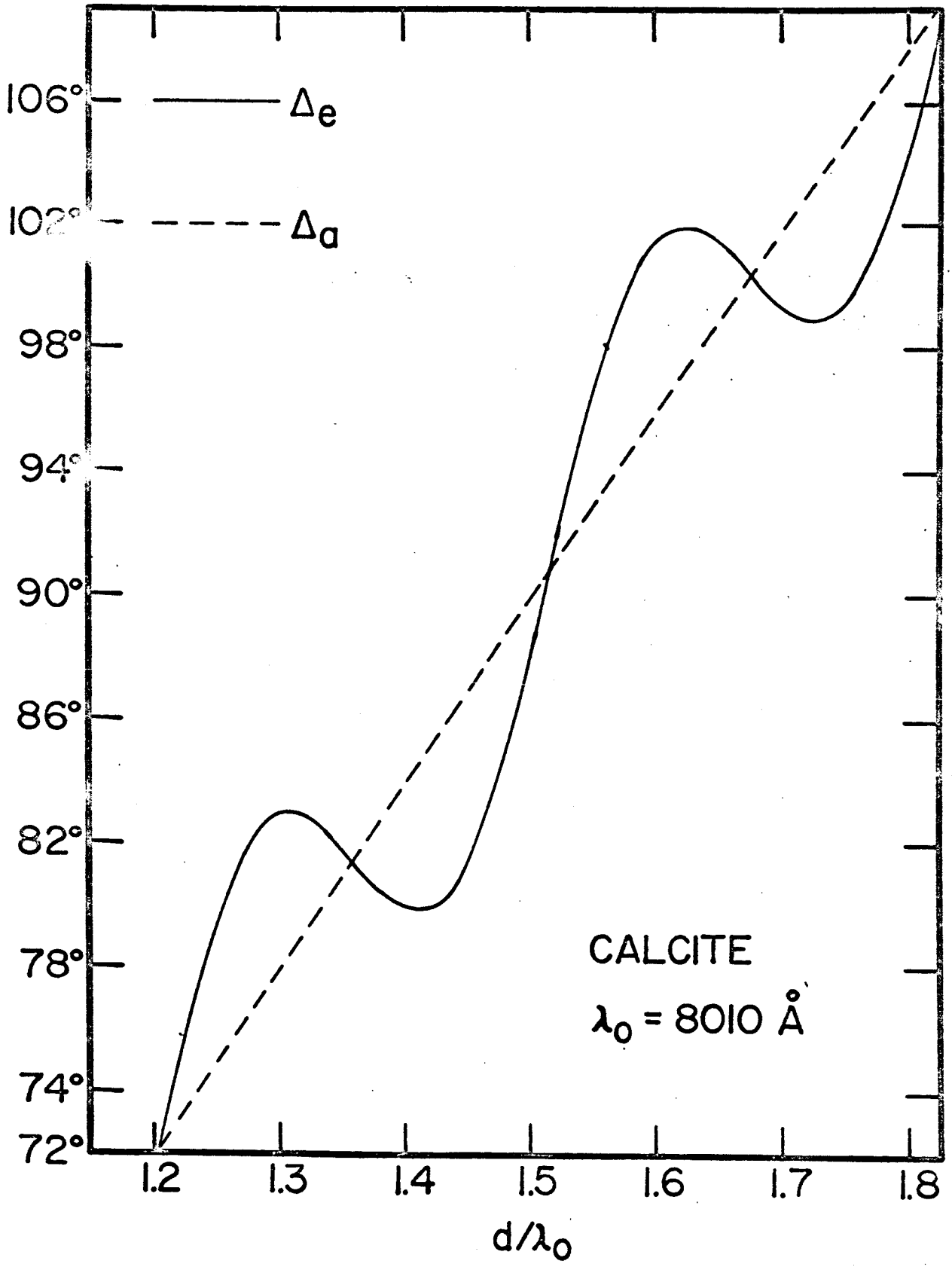

106°
102°
98°
94°
90°
86°
82°
78°
74°
72°
Δ_e
Δ_a
CALCITE
λ_0 = 8010 Å
1.2
1.3
1.4
1.5
1.6
1.7
1.8
d/λ_0

Fig. B-5

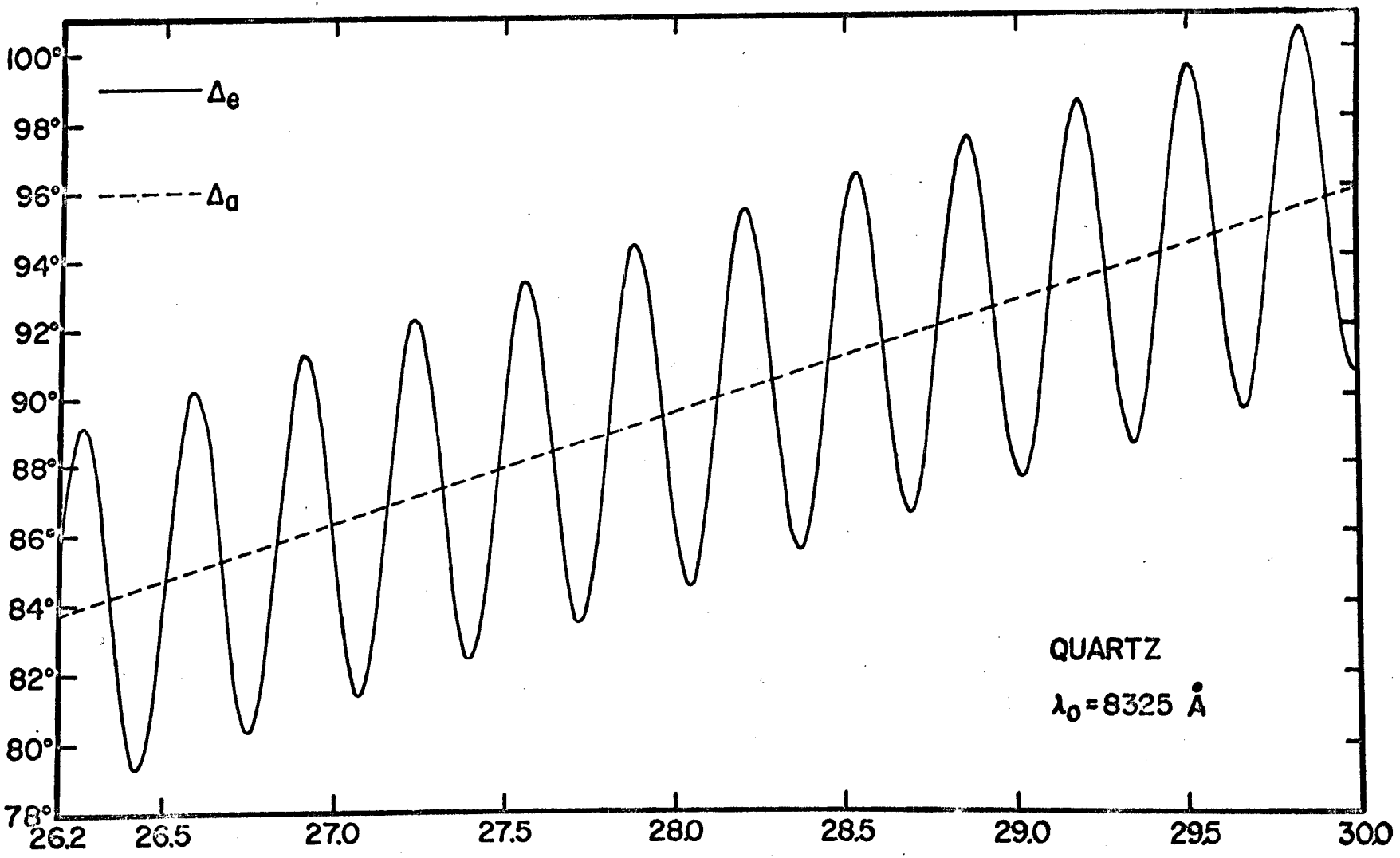

100°
98°
96°
94°
92°
90°
88°
86°
84°
82°
80°
78°
Δ_e
Δ_a
QUARTZ
λ_0 = 8325 Å
26.2
26.5
27.0
27.5
28.0
28.5
29.0
29.5
30.0

Fig. B-6

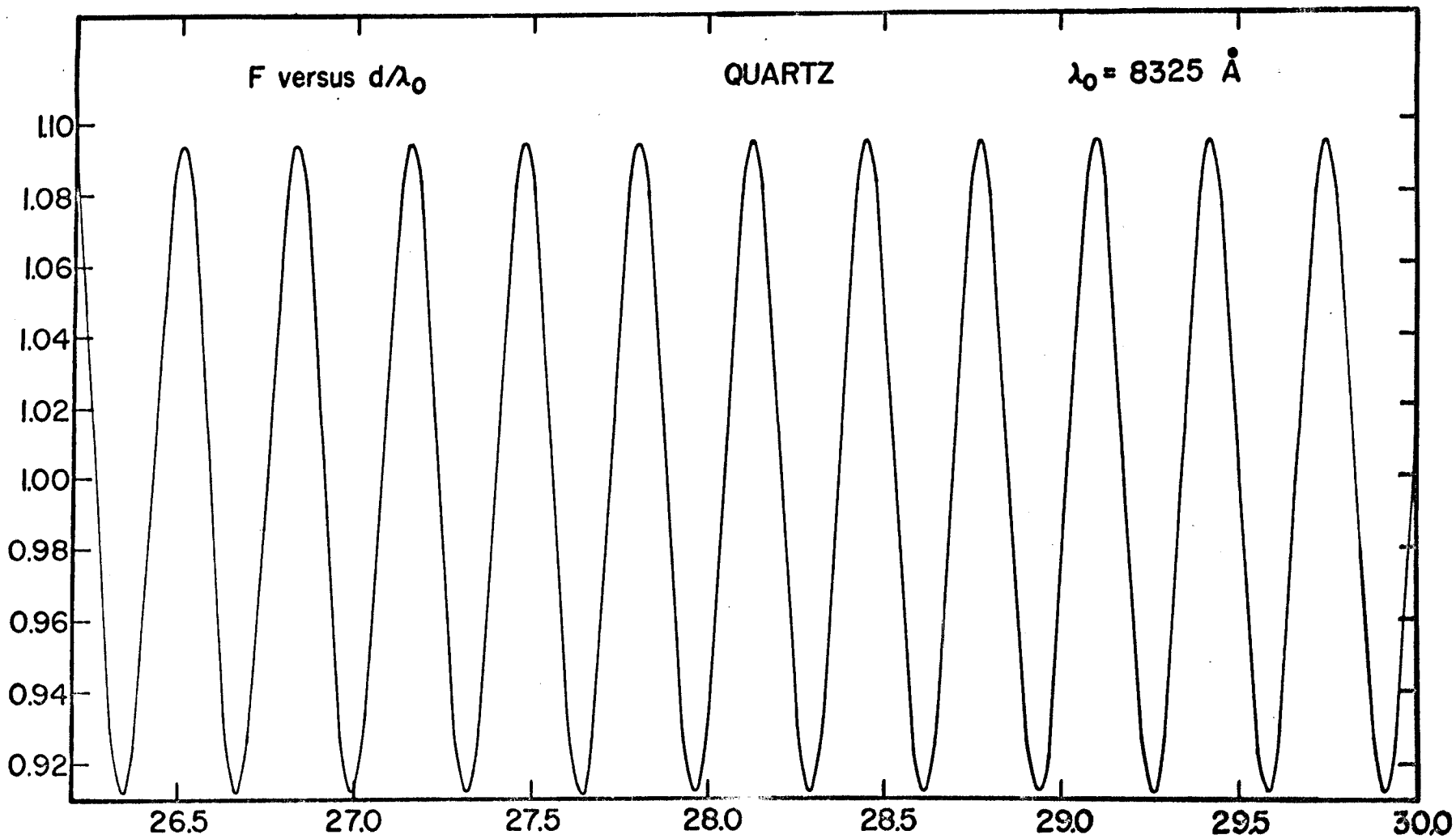

F versus d/λ0
QUARTZ
λ0 = 8325 Å
1.10
1.08
1.06
1.04
1.02
1.00
0.98
0.96
0.94
0.92
26.5
27.0
27.5
28.0
28.5
29.0
29.5
30.0

Fig. B-7

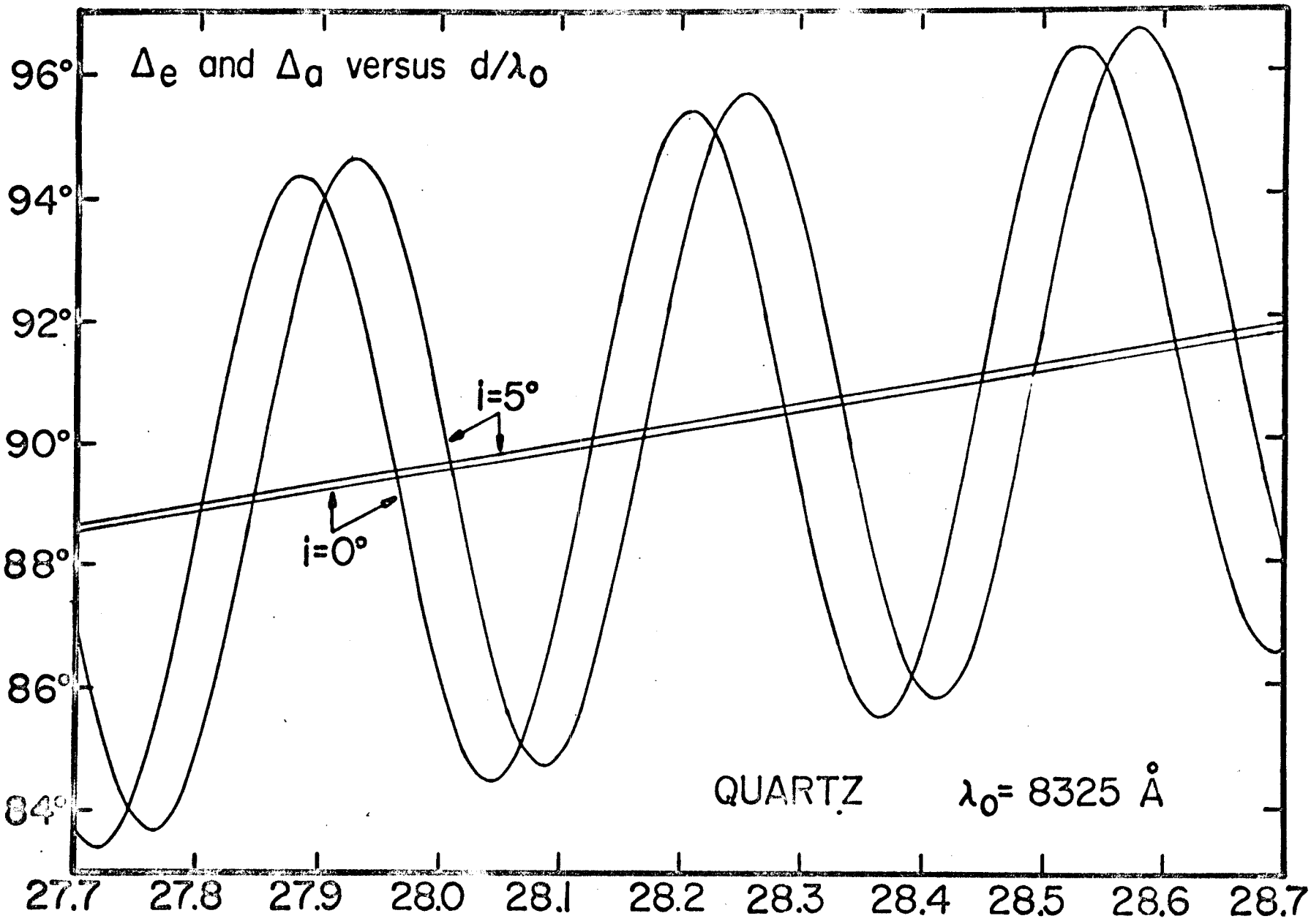

Δ_e and Δ_o versus d/λ_o
96°
94°
92°
90°
88°
86°
84°
i=5°
i=0°
QUARTZ
λ_o= 8325 Å
27.7
27.8
27.9
28.0
28.1
28.2
28.3
28.4
28.5
28.6
28.7

Fig. B-8

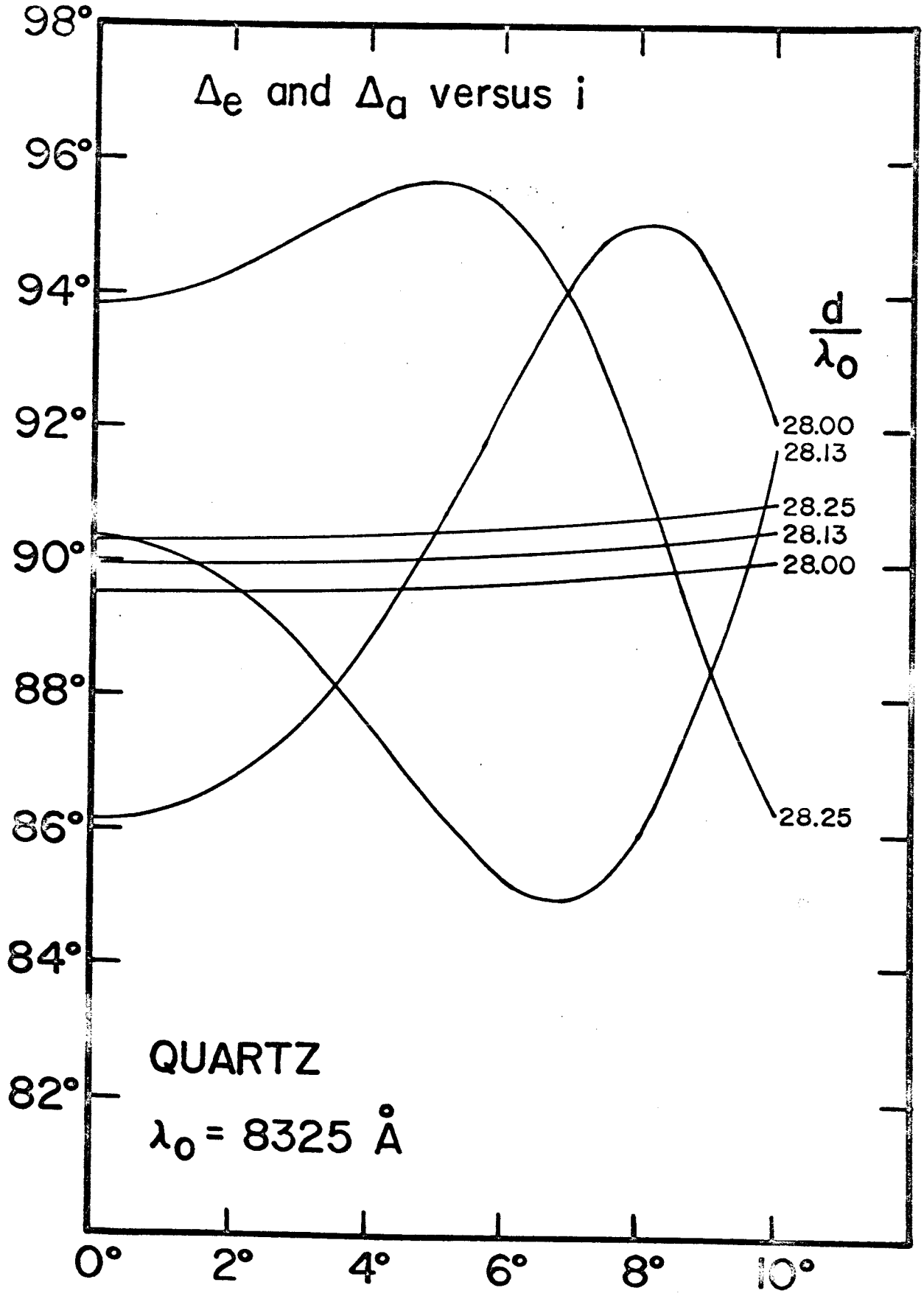

98°
96°
94°
92°
90°
88°
86°
84°
82°
Δe and Δa versus i
d/λo
28.00
28.13
28.25
28.13
28.00
28.25
QUARTZ
λo = 8325 Å
0°
2°
4°
6°
8°
10°

Fig. B-9

APPENDIX C

Wave Optics Theory of Rotary Compensators

This appendix is a relatively self-contained work and has been published in the Journal of the Optical Society of America, vol. 54, p. 1340 (1964). The differences between this appendix and the published article are only minor. Permission to reproduce the published work has been granted by the Journal of the Optical Society of America. The figures and figure captions relevant to this appendix can be found in numerical sequence at the end of the appendix.

Wave Optics Theory of Rotary Compensators*

D. A. Holmes
Electrical Engineering Department
Carnegie Institute of Technology
Pittsburgh 13, Pennsylvania

Abstract

The theory conventionally used to describe the operation of Berek and Ehringhaus rotary compensators is based on geometrical optics and, hence, is not exact. When rotary compensators are analyzed within the framework of classical electromagnetic theory, exact solutions for the phase difference and amplitude ratio of the transmitted light can be determined. The approximate and exact solutions differ in some interesting ways which become of substantial importance in the examination of monochromatic plane waves of light. In particular the discrepancies between exact and approximate solutions become more pronounced at high angles of incidence. Exact theory predicts the possibility of using a high refractive index isotropic plate for measuring small phase differences at relatively long wavelengths.

---

*This work was supported by the National Aeronautics and Space Administration under contract NAS8-5269 and forms a portion of a thesis submitted to Carnegie Institute of Technology in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

I. INTRODUCTION

Rotary compensators are often used in determining the character of elliptically polarized, monochromatic, plane waves of light. We can obtain a qualitative understanding of the operation of rotary compensators by examining the geometry depicted in Fig. C-1. Crystalline plates 1 and 2 are mounted back to back, and an xyz coordinate system is defined such that the xy plane is at the interface of plate 1 and medium 1. For convenience, the orientations of the crystalline plates are restricted to the extent that the principal dielectric axes of each plate are assumed to be parallel to the x, y and z coordinate axes. The subscripted n quantities are then the principal refractive indices of plates 1 and 2. The surrounding isotropic media I and II have unity index of refraction. The plane wave under examination is incident such that the plane of incidence is the xz plane while the incident wave normal makes an angle i with respect to the z axis. The electric field vector of the incident wave can be decomposed into two components; one component is polarized parallel to the plane of incidence and is designated the p component, while the other is polarized perpendicular to the plane of incidence and is designated the s component. This classification is similar to that commonly used in ellipsometry work.[1,2] The ratio of the p and s components of the incident wave can be characterized by an amplitude ratio $\tan\alpha_i$ and a phase difference δ_i as follows

$$E_{pi}/E_{si} = e^{j\delta_i} \tan\alpha_i , \tag{1}$$

---

[1] R. J. Archer, J. Opt. Soc. Am. 52, 970 (1962).

[2] F. Partovi, J. Opt. Soc. Am. 52, 918 (1962).

where E_{pi} and E_{si} are the respective p and s components of the incident electric field vector.

After the incident wave propagates through the system of birefringent plates it emerges into medium II as a plane wave whose p and s electric field components can be written as

$$E_{po}/E_{so} = e^{j\delta_o} \tan\alpha_o . \qquad (2)$$

The components of the output wave are related to the components of the input wave by

$$e^{j\delta_o} \tan\alpha_o = e^{j(\Delta+\delta_i)} T \tan\alpha_i . \qquad (3)$$

In (3), Δ is the phase difference introduced by transmission through the plates and T is an amplitude ratio factor. The quantities Δ and T are considered to be known functions of the principal indices of refraction of the plates, the plate thicknesses d_1 and d_2, the angle of incidence, and the wavelength.

The optical measurement of α_i and δ_i is then accomplished by rotating the plates about an axis parallel to the y axis, thus changing the angle of incidence, until the phase difference Δ is adjusted to such a value that the output wave is linearly polarized. The linearly polarized output wave can then be extinguished by an analyzer. From the rotational setting of the birefringent plates we know Δ and hence can calculate δ_i. From the analyzer setting α_o is determined and, since T is known from the rotational setting of the plates, we can then calculate α_i. Although the complete polarization state[3] of the incident light can not be determined by the above measurement, in many cases, particularly in ellipsometry experiments, the determination of α_i and δ_i is sufficient.

---

[3] C. A. Skinner, J. Opt. Soc. Am. and Rev. Sci. Instr. 10, 491 (1925).

A system of birefringent plates that is used for compensating phase differences in the manner described above is called a rotary compensator. This definition of rotary compensators then excludes those of the Senarmont type[4,5], which are usually provided with a mechanism to achieve rotation about an axis parallel to, rather than perpendicular to, the direction of propagation of the incident wave.

Among the better known of the present day commercially available rotary compensators are those of the Berek type[6-8], manufactured by E. Leitz, Inc., and those of the Ehringhaus type[9], manufactured by Carl Zeiss, Inc. The typical Berek compensator consists of a single calcite plate (uniaxial crystal) usually about 0.1 mm thick, with the optic axis normal to the plate surfaces. The geometry of Fig. C-1 corresponds to that of a calcite Berek compensator if we set

$$\begin{aligned} n_{x1} = n_{x2} = n_{y1} = n_{y2} &= \omega , \\ n_{z1} = n_{z2} &= \epsilon , \\ d_1 + d_2 = d &\cong 0.1\ \text{mm} , \end{aligned} \tag{4}$$

where ω is the ordinary refractive index and ϵ is the extraordinary refractive index. A typical Ehringhaus compensator uses two plates of quartz (uniaxial crystal). The plates are of equal thickness, usually 1.0 mm each. The optic axes of the two plates are mutually perpendicular and are both parallel to the plate surfaces. One of the optic axes is parallel to the axis of rotation. The geometry of

---

[4] H. G. Jerrard, J. Opt. Soc. Am. 38, 35 (1948).

[5] A. C. Hall, J. Opt. Soc. Am. 53, 801 (1963).

[6] M. Berek, Mikroskopische Mineralbestimmung mit Hilfe der Universaldrehtischmethoden (Gebr. Borntrager, Berlin, 1924).

[7] F. Rinne und M. Berek, Anleitung zu Optischen Untersuchugen mit dem Polarizationsmikroskop (Schweizerbart'sche Verlagsbuchhandlung, Stuttgart, 1953).

[8] M. Berek, Zentralblatt f. Mineralogie 388, 427, 464, 580 (1913).

[9] A. Ehringhaus, Z. Kristallogr, 76, 315 (1931); 98, 394 (1938); 102, 85 (1939).

Fig.C-1 corresponds to that of a quartz Ehringhaus compensator if we set

$$\begin{aligned} n_{x1} = n_{z1} = n_{y2} = n_{z2} &= \omega, \\ n_{y1} = n_{x2} &= \epsilon, \\ d_1 = d_2 = d &\cong 1.0\ \text{mm}. \end{aligned} \tag{5}$$

In order to predict the behavior of a rotary compensator it is required that one know how Δ and T vary as a function of the angle of incidence for a given set of plates and a given wavelength. In the past, rotary compensators have been analyzed by using the tools of geometrical optics. Burri[10] has shown that the geometrical optics equations for the phase difference can be written as

$$\Delta_a = \frac{2\pi d}{\lambda_0}\left\{(\omega^2 - \sin^2 i)^{1/2} - \frac{\omega}{\epsilon}\cdot(\epsilon^2 - \sin^2 i)^{1/2}\right\}, \tag{6}$$

for the calcite Berek compensator and

$$\Delta_a = \frac{2\pi d}{\lambda_0}\left\{(\epsilon^2 - \sin^2 i)^{1/2} - \frac{\epsilon}{\omega}(\omega^2 - \sin^2 i)^{1/2}\right\}, \tag{7}$$

for the quartz Ehringhaus compensator, where λ_0 is the vacuum wavelength. In the geometrical optics analysis T is set equal to unity.

The above solutions for the phase difference, however, are only approximate (hence the subscript a attached to Δ) because the geometrical optics approach ignores the existence of multiple internal reflections between the surfaces of the crystalline plates. Recent experimental[11] and theoretical[12] work has shown that resonance effects can be very important when wave plates and Senarmont compensators are used, in an air environment, for examining elliptically polarized, mono-

---

[10] Conrad Burri, Z. Angw. Math. Phys. 4, 418 (1953).

[11] H. Weinberger and J. Harris, J. Opt. Soc. Am. 54, 552 (1964).

[12] D. A. Holmes, J. Opt. Soc. Am. 54, 1115 (1964), (App. B).

chromatic, plane waves. From information I have received through several private communications with the manufacturers of Berek and Ehringhaus compensators, it is apparent that equations (6) and (7) are currently being used in the calibration of calcite Berek and quartz Ehringhaus rotary compensators. Also, Gahm[13] has recently presented a study of error sources and measurement accuracy in the use of quartz Ehringhaus compensators. Gahm, however, used the geometrical optics approach and therefore overlooked those errors which may arise because of interference or resonance effects. In the present work I will show that resonance effects are quite significant when rotary compensators in an air environment are used to examine elliptically polarized, monochromatic, plane waves.

In section 2 the formal Maxwell equation solution is given for the problem of transmission through two homogeneous, non-absorbing, non-optically active, crystalline plates mounted back to back and immersed in air. The generality of the anisotropy is limited to that shown in Fig.C-1. The geometrical optics solutions are then obtained from the exact or wave optics solutions under special conditions. Some useful approximations are also developed. In sections 3 and 4, comparisons between the wave optics and geometrical optics solutions are given for a calcite Berek compensator and a quartz Ehringhaus compensator, respectively. The geometries chosen for numerical illustration correspond to those of the commercially available devices. In section 5, an isotropic rotary retardation plate is discussed.

2. ANALYSIS

In this section we shall consider the electromagnetic solution to the problem of transmission through two biaxial crystalline plates corresponding to the geometry of Fig.C-1. The objective is to determine

---

[13] J. Gahm, Zeiss Mitteilungen uber Fortschritte der Technischen Optik 3, 152 (1964).

the expressions for the phase difference and the amplitude ratio factor introduced in equation (3).

2.1 Wave Optics Solutions

The reflection and transmission properties of birefringent crystals have been studied in the past. Winterbottom[14], for example, has summarized the wave optics solutions for reflection from a semi-infinite biaxial crystal and from a semi-infinite uniaxial crystal covered with a uniaxial film. Schopper[15] has given a rather comprehensive treatment of transmission through a single biaxial absorbing crystalline plate. The wave optics approach employed by Schopper can be used to determine the transmission characteristics of the two lossless biaxial plates shown in Fig.C-1. From a wave optics analysis we then find that the phase difference and the amplitude ratio factor are given by

$$
\begin{aligned}
T e^{j\Delta_e} = & \left\{ \cos\beta_{s1} d_1 \cos\beta_{s2} d_2 - K_{s12} \sin\beta_{s1} d_1 \sin\beta_{s2} d_2 \right. \\
& \left. + j\left[K_{s1} \sin\beta_{s1} d_1 \cos\beta_{s2} d_2 + K_{s2} \cos\beta_{s1} d_1 \sin\beta_{s2} d_2 \right] \right\} \\
& \times \left\{ \cos\beta_{p1} d_1 \cos\beta_{p2} d_2 - K_{p12} \sin\beta_{p1} d_1 \sin\beta_{p2} d_2 \right. \\
& \left. + j\left[K_{p1} \sin\beta_{p1} d_1 \cos\beta_{p2} d_2 + K_{p2} \cos\beta_{p1} d_1 \sin\beta_{p2} d_2 \right] \right\}^{-1},
\end{aligned}
\tag{8}
$$

where

$$
\beta_{sk} = (2\pi/\lambda_0)\left(n_{yk}^2 - \sin^2 i\right)^{1/2}, \quad k = 1, 2,
\tag{9}
$$

---

[14] A. B. Winterbottom, Det KGL Norske Vindenskabers Selskabs Skrifter 1, 27, 37 (1955).

[15] H. Schopper, Z. Phys. 132, 146 (1952).

$$\beta_{pk} = (2\pi n_{xk}/\lambda_0 n_{zk})(n_{zk}^2 - \sin^2 i)^{1/2}, \quad k = 1,2, \tag{10}$$

$$2K_{pk} = \frac{n_{xk} n_{zk} \cos i}{(n_{zk}^2 - \sin^2 i)^{1/2}} + \frac{(n_{zk}^2 - \sin^2 i)^{1/2}}{n_{xk} n_{zk} \cos i}, \quad k = 1,2, \tag{11}$$

$$2K_{sk} = \frac{\cos i}{(n_{yk}^2 - \sin^2 i)^{1/2}} + \frac{(n_{yk}^2 - \sin^2 i)^{1/2}}{\cos i}, \quad k = 1,2, \tag{12}$$

$$2K_{p12} = \frac{n_{x2} n_{z2}}{n_{x1} n_{z1}} \cdot \left(\frac{n_{z1}^2 - \sin^2 i}{n_{z2}^2 - \sin^2 i}\right)^{1/2} + \frac{n_{x1} n_{z1}}{n_{x2} n_{z2}} \cdot \left(\frac{n_{z2}^2 - \sin^2 i}{n_{z1}^2 - \sin^2 i}\right)^{1/2}, \tag{13}$$

$$2K_{s12} = \left(\frac{n_{y1}^2 - \sin^2 i}{n_{y2}^2 - \sin^2 i}\right)^{1/2} + \left(\frac{n_{y2}^2 - \sin^2 i}{n_{y1}^2 - \sin^2 i}\right)^{1/2}. \tag{14}$$

The subscript e has been affixed to Δ to indicate that the phase difference Δ_e is calculated by exact theory.

2.2 Geometrical Optics Solutions

The geometrical optics solutions can be conveniently obtained from the wave optics solutions by ignoring the multiple reflection effects. The resonance effects are eliminated by setting $K_{s12} = K_{p12} = K_{s1} = K_{s2} = K_{p1} = K_{p2} = 1$. By replacing Δ_e in (8) by Δ_a we then have

$$T = 1, \tag{15}$$

$$\Delta_a = (\beta_{s1} - \beta_{p1})\, d_1 + (\beta_{s2} - \beta_{p2})\, d_2. \tag{16}$$

When the principal indices of refraction and the plate thicknesses are specified as in (4), (16) reduces to (6). Similarly, when conditions (5) are used, (16) reduces to (7).

2.3 Approximations

From (8) we may write

$$\Delta_e = \arctan\left[\frac{K_{s1}\tan\beta_{s1}d_1 + K_{s2}\tan\beta_{s2}d_2}{1 - K_{s12}\tan\beta_{s1}d_1\tan\beta_{s2}d_2}\right]$$

$$- \arctan\left[\frac{K_{p1}\tan\beta_{p1}d_1 + K_{p2}\tan\beta_{p2}d_2}{1 - K_{p12}\tan\beta_{p1}d_1\tan\beta_{p2}d_2}\right] . \qquad (17)$$

Now let us assume that the numerical values of the respective principal indices of refraction for crystals 1 and 2 allow the following approximations to be made.

$$\begin{aligned} &K_{s12} \cong 1 \quad , \quad K_{p12} \cong 1 \, , \\ &\left| K_{s1} - K_{s2} \right| \ll K_{s1} + K_{s2} \, , \\ &\left| K_{p1} - K_{p2} \right| \ll K_{p1} + K_{p2} \, . \end{aligned} \qquad (18)$$

Using (18) in (17) we then find that

$$\Delta_e \cong \arctan\left[\tfrac{1}{2}(K_{s1} + K_{s2})\tan(\beta_{s1}d_1 + \beta_{s2}d_2)\right]$$

$$- \arctan\left[\tfrac{1}{2}(K_{p1} + K_{p2})\tan(\beta_{p1}d_1 + \beta_{p2}d_2)\right] . \qquad (19)$$

Equation (19) is now in such a form that we may use an approximation developed in a previous work on the wave optics theory of retardation plates[12] to obtain

$$\Delta_e \cong \Delta_a + \Phi \cos\theta \sin\Delta_a , \tag{20}$$

where $$\Phi = (K_{s1} + K_{s2} + K_{p1} + K_{p2} - 4)/4 ,$$

$$\theta = (\beta_{s1} + \beta_{p1}) d_1 + (\beta_{s2} + \beta_{p2}) d_2 .$$

The restrictions on (20) are

$$(K_{s1} + K_{s2} - 2)/2 \ll 1 , \; (K_{p1} + K_{p2} - 2)/2 \ll 1 . \tag{21}$$

As the angle of incidence i is varied we note that $\cos\theta$ is a much more rapidly varying function of i than is $\sin\Delta_a$. In later sections we shall refer to the function defined by $\pm \Phi \sin\Delta_a$ as the envelope of $\Delta_e - \Delta_a$ or simply as the envelope function.

3. CALCITE BEREK COMPENSATOR

By substituting (4) into (8) we have that

$$\Delta_e = \arctan\left[\frac{K_s \tan\beta_s d - K_p \tan\beta_p d}{1 + K_s K_p \tan\beta_s d \tan\beta_p d}\right] , \tag{22}$$

$$T = \left[\frac{1 + (K_s^2 - 1)\sin^2\beta_s d}{1 + (K_p^2 - 1)\sin^2\beta_p d}\right]^{1/2} , \tag{23}$$

where

$$\beta_s = (2\pi/\lambda_0)(\omega^2 - \sin^2 i)^{1/2} \quad , \beta_p = (2\pi\omega/\epsilon\lambda_0)(\epsilon^2 - \sin^2 i)^{1/2} ,$$

$$2K_s = \frac{\cos i}{(\omega^2 - \sin^2 i)^{1/2}} + \frac{(\omega^2 - \sin^2 i)^{1/2}}{\cos i} ,$$

$$2K_p = \frac{\epsilon\omega\cos i}{(\epsilon^2 - \sin^2 i)^{1/2}} + \frac{(\epsilon^2 - \sin^2 i)^{1/2}}{\epsilon\omega\cos i} .$$

The geometrical optics value of the phase difference is given by $\Delta_a = (\beta_s - \beta_p)\, d$, which reduces to (6).

In Fig. C-2a, b is shown the variation of $\Delta_e - \Delta_a$ as a function of angle of incidence i for a wavelength of 8010 Angstroms. The optical constants[16] for calcite are taken as ω = 1.64869 and ϵ = 1.48216. At this wavelength, we have that $d = 120\,\lambda_o = 0.09612$ mm. This thickness value closely approximates the 0.1 mm thickness of the calcite plate in the commercially available Berek compensator. The range of $0^\circ \leqq i \leqq 38^\circ$ covers about $3\frac{1}{2}$ orders of Δ_a . The envelope function, given by $\pm \Phi \sin \Delta_a$ is also shown. As an initial observation we note that the envelope function follows fairly closely the peaks of $\Delta_e - \Delta_a$ and also defines the "recurrent" variation of $\Delta_e - \Delta_a$ with i . As the angle of incidence is increased the zero crossings of the envelope function become closer spaced while the maximum values increase.

We note that Δ_e can differ from Δ_a by as much as 7.4° at the higher values of i . We have also plotted calculations for a thin slab, or $d = 50\,\lambda_o$, over the range $0^\circ \leqq i \leqq 19^\circ$. In general, the number of zero crossings of $\Delta_e - \Delta_a$ over a given range of i increases with the thickness of the slab.

Suppose now we consider the effects introduced by using a divergent beam of light, assuming that, for very small angles of divergence, say 0.2°, or less, the beam can be considered as a super-position of plane waves traveling in slightly different directions. For purposes of discussion we shall confine the beam divergence to the xz plane. Suppose further that the input beam is elliptically polarized and that compensation is attempted in one of the higher orders, say at $i = 26^\circ$.

---

[16] For illustrative numerical examples we have chosen optical constants for calcite and quartz corresponding to near infrared wavelengths because we are using gallium arsenide injection laser sources (8400Å) in some of our work. We assume that no absorption occurs at the wavelengths used. The general concepts advanced in this work, however, clearly apply at other wavelengths. The optical constants were taken from D. E. Gray, coordinating editor, American Institute of Physics Handbook (McGraw-Hill Book Co., Inc., New York, 1963), 2nd edition, Calcite, p. 6-18; Crystal Quartz, p. 6-33.

If the beam divergence is contained in $25.9^{\circ} \leqq i \leqq 26.1^{\circ}$, a spread of 18° in Δ_e is obtained and perfect compensation of the entire beam is impossible. This would result in failure to obtain extinction with an analyzer. The inability to find extinction is predicted to a somewhat lesser extent by approximate theory, the spread in Δ_a being about 9° for the beam divergence given above.

In Fig. C-3a, b, T is shown as a function of angle of incidence. We note that T is roughly periodic with Δ_a and can be as large as 1.2 in the fourth order of Δ_a. An interesting feature is that the minima of T tend to rise toward unity as i increases while the maxima of T tend to increase away from unity as i increases. Suppose now that the compensator is in the diagonal position, i.e., $|E_{pi}/E_{si}| = \tan\alpha_i = 1$. If the incident wave is compensated, the geometrical optics approximation predicts that the resultant linear vibration will make an angle of 45° with respect to the y axis. Suppose now, that $T = 1.15$, for the rotational setting at which compensation occurs. Then according to exact theory the linear vibration is at $\arctan(1.15) \cong 49^{\circ}$, a discrepancy of 4° with the approximate theory.

The errors inherent in the geometric approach are more serious at high angles of incidence. For example, at angles of incidence close to 60°, the maximum excursion of $\Delta_e - \Delta_a$ from zero is about 13° while T can be as large as 1.58.

4. QUARTZ EHRINGHAUS COMPENSATOR

By substituting (5) into (8), T and Δ_e for the quartz Ehringhaus compensator can be found. For this compensator, (8) and (17) do not significantly simplify, therefore the explicit relations for T and Δ_e are not contained in this section.

At a wavelength $\lambda_o = 8325\text{Å}$, the optical constants[16] for quartz are taken as $\epsilon = 1.54661$ and $\omega = 1.53773$. In Fig. C-4, the envelope of $\Delta_e - \Delta_a$ is plotted for an Ehringhaus compensator with two crystalline quartz plates, each plate of thickness $d = 1200\lambda_o = 0.999$ mm. This

thickness value is virtually the same as that found in the commercially available compensator. The envelope function for the Ehringhaus compensator exhibits the same general characteristics as does the envelope function for the Berek compensator. The fine structure in $\Delta_e - \Delta_a$ which is controlled by the $\cos\theta$ term in the approximation (20) is, however, rather different for the respective compensators. The ranges in i delineated by the dotted lines in Fig. C-4 are shown on an expanded scale in Figs. C-5 and 6 to illustrate the implications of the resonance effects in quartz Ehringhaus compensators.

The actual curves of Δ_e and Δ_a are shown in Fig. C-5 for the range $5.0^\circ \leqq i \leqq 7.5^\circ$. The range of Δ_a covered is only about 15° with the maximum value of Δ_a being about 28°, yet requirements on beam collimation are already becoming stringent. For example, consider a diverging beam which is confined to the range $6.72^\circ \leqq i \leqq 6.80^\circ$. Between these limits there is a spread in Δ_e of 4.45° while the spread in Δ_a is 0.53°. Also, many different rotational settings will yield the same value of phase difference according to exact theory. If, for example, a perfectly collimated incident wave can be characterized by a phase difference $\delta_i = -24^\circ$, we see from Fig. 5 that 9 different positions of the compensator could be used to compensate the incident wave.

For the range $47.4^\circ \leqq i \leqq 47.6^\circ$ we have plotted $\Delta_e - \Delta_a$ in Fig. C-6. Over this small range the envelope function, also shown, is practically a straight line. Let us consider the curve for $d = 1200\lambda_o$ and assume that an incident divergent beam is contained in the interval $47.480^\circ \leqq i \leqq 47.495^\circ$. The spread in Δ_e over this interval is 12.58°, while the spread in Δ_a is only 0.32°. In practice, a beam divergence of $0.015^\circ \cong 0.00026$ radian is quite difficult to obtain, except, perhaps, from a gas laser. Another interesting feature is the accuracy with which the rotational setting can be determined. For numerical purposes let us suppose the setting can be measured to an accuracy of $1' \cong 0.017^\circ$. We see that the setting accuracy is not sufficient to resolve the variation of Δ_e in the range of incidence angles considered in Fig. C-6. The dependence of Δ_e upon d/λ_o is also shown in Fig. C-6. For a change

in d/λ_o of 1/10 the corresponding changes in Δ_e are quite significant while the corresponding changes in Δ_a are negligible. Exact theory, in contrast to approximate theory, prescribes very stringent conditions on the monochromaticity of the incident wave and on the surface flatness of the quartz plates.

5. ISOTROPIC PLATE

The properties of an homogeneous, isotropic dielectric plate immersed in air are of particular interest in optics[17]. In this section we wish to discuss the"compensator" properties of this geometry. In Fig. C-1 we set all of the refractive indices equal to n and $d_1 + d_2 = d$. For this case, from (8) we obtain

$$\Delta = \tan^{-1}\left[\frac{(K_s - K_p)\tan\beta d}{1 + K_s K_p \tan^2\beta d}\right], \tag{24}$$

$$T = \left[\frac{1 + K_s^2 \tan^2\beta d}{1 + K_p^2 \tan^2\beta d}\right]^{1/2}, \tag{25}$$

where $$\beta d = (2\pi d/\lambda_o)(n^2 - \sin^2 i)^{1/2},$$

$$2K_s = \frac{\cos i}{(n^2 - \sin^2 i)^{1/2}} + \frac{(n^2 - \sin^2 i)^{1/2}}{\cos i},$$

$$2K_p = \frac{n^2 \cos i}{(n^2 - \sin^2 i)^{1/2}} + \frac{(n^2 - \sin^2 i)^{1/2}}{n^2 \cos i}.$$

---

[17] Some aspects of propagation through an isotropic slab are discussed by M. Born and E. Wolf, Principles of Optics (Pergamon Press, New York, 1959), pp. 60 et seq.

In (24), Δ is the value of phase shift calculated by wave optics theory. Geometrical optics predicts that the phase difference produced by transmission through an isotropic plate at oblique incidence will be zero.

From inspection of (24) and (25) we note that Δ and T are periodic functions of the plate thickness d. In particular, Δ is a continuous function of d and passes through finite maxima and minima. We shall now determine how these maximum values of Δ depend upon i and n. If we consider, for the moment, that i and n are fixed and then maximize Δ with respect to d, we find that the maximum values of Δ are found for thickness values which satisfy $\tan^2 \beta d = (K_s K_p)^{-1}$. The corresponding maxima in Δ are then given by

$$\Delta_{max} = \tan^{-1}\left\{(K_s/4K_p)^{1/2} - (K_p/4K_s)^{1/2}\right\} . \qquad (27)$$

Equation (27) represents the maximum possible phase difference for a given dielectric material at a fixed angle of incidence. The quantity Δ_{max} is plotted as a function of angle of incidence with refractive index as a parameter in Fig. C-7. It should be noticed that, for each value of Δ_{max} shown in Fig. C-7, the corresponding optimum value of d is, in general, different. The maximum phase difference Δ_{max} increases as either i or n is increased, however, Δ_{max} will never reach 90° for a finite value of n. It should be pointed out that Δ given by (24) approaches zero as i approaches 90°, except when $\tan^2 \beta d = (K_p K_s)^{-1}$.

As Bergman[18] has pointed out, a compensator of phase difference Δ can be used to compensate elliptically polarized light provided that

$$(b/a) \leqq \tan(\Delta/2) \leqq (a/b) , \qquad (19)$$

---

[18] D. Bergman, J. Opt. Soc. Am. 52, 1080 (1962).

where b/a is the ratio of the minor axis to the major axis of the ellipse. From Fig. C-7 we see that it is theoretically possible to use an isotropic plate to compensate a monochromatic plane wave of elliptically polarized light. The physical and optical properties of the plate would limit the range of ellipticities that could be measured. In particular, circularly polarized light cannot be compensated with a single slab.

Let us consider some of the important details concerning the design of a rotary isotropic slab retardation plate. It is clear from Fig.C-7 that a high refractive index material is desired in order to obtain a given value of Δ without having to go to extreme angles of incidence. Another important consideration is the variation of Δ with i. From (24) we note that Δ passes through zero when $\beta d = m\pi/2$, where m is an integer. For maximum sensitivity it would be desirable to have the number of zero crossings of Δ be as small as possible over the working range $i_1 \leq i \leq i_2$. The number of zero crossings in the working range can be determined by counting the values of m which satisfy

$$(4d/\lambda_0)(n^2 - \sin^2 i_2)^{1/2} \leq m \leq (4d/\lambda_0)(n^2 - \sin^2 i_1)^{1/2}. \tag{28}$$

For a given value of n, the number of zero crossings of Δ in the range $i_1 \leq i \leq i_2$ increases as d/λ_0 increases. In order to keep d/λ_0 as small as possible it becomes necessary to use very thin crystal plates or, alternatively, long wavelength radiation. For physically realizable values of d one may find that a given sensitivity requirement may only be satisfied for wavelengths in the infrared.

With the above considerations in mind let us consider, as a numerical example, the phase difference introduced by transmission through a thallium bromide iodide crystalline plate at a wavelength of $\lambda_0 = 10\mu$. We assume that absorption by the crystal is negli-

gible at this wavelength[19]. In Fig. C-8 is shown Δ versus i for this hypothetical situation. For both plate thicknesses, Δ has eight zero crossings in the range $0° \leqq i \leqq 90°$. This can be verified from (28). It is seen that Δ has a strong dependence on d. For this particular plate we see that the maximum ellipticity which could be measured is about 1/5, if we restrict the measuring range to $0° \leqq i \leqq 60°$. Thus, it is seen that a single plate would serve primarily as a device for measuring small retardations or for extending the frequency range of a quarter-wave plate. By using several plates in series one could measure larger ellipticities. For this latter situation one must be careful to space the plates sufficiently far apart so that reflections from a given plate cannot be intercepted by the previous plate.

In analyzing the transmitted light one must be aware that the direction of the resultant linear vibration is affected by the factor T, given in (25). Note that the amplitude ratio factor T will always be greater than or equal to unity because $K_s > K_p$. An interesting feature of the isotropic slab retardation plate is that it can be rendered essentially inoperative in an optical system by adjusting the angle of incidence such that $\tan \beta d = 0$. In this case $T = 1$ and $\Delta = 0°$. At the zero setting, or $i = 0°$, we also have $T = 1$ and $\Delta = 0°$, however, the intensity of the transmitted light is diminished due to reflections back into the optical system. Because of the relatively large dielectric constants required for the plate, the reflections at normal incidence would not be insignificant and, as discussed by Winterbottom[20], would be a source of error in an optical measurement.

---

[19] A summary of the optical properties of thallium bromide iodide (KRS-5) is given in "Synthetic Optical Crystals", (The Harshaw Chemical Company, Cleveland 6, Ohio, 1955) rev. ed., pp. 23-24. We have taken the index of refraction at 10μ as n = 2.37274.

[20] Reference 14, p. 68.

Very pure isotropic semiconductor crystals used at wavelengths longer than that corresponding to the fundamental absorption edge also have potential as isotropic slab retardation plates. For example, with a silicon crystal having a refractive index of approximately 3.5, one could expect a maximum phase difference of about 35° at an angle of incidence of 65°.

6. SUMMARY AND DISCUSSION

The phase difference Δ_e and the amplitude ratio factor T have been given for the problem of transmission through two biaxial plates at an oblique angle of incidence. The orientations of the plates were restricted to the case for which one principal axis of each plate is perpendicular to the plane of incidence. The solutions for Δ_e and T can be used to predict the behavior of other optical devices utilizing the geometry of Fig.C-1. The Soleil compensator[21], for example, is effectively a two plate compensator which is used at normal incidence, i.e., $i = 0^\circ$. By substituting $i = 0^\circ$ and say, $n_{x1} = n_{y2} = \epsilon$ and $n_{x2} = n_{y1} = \omega$, into (8), one could predict the exact theory behavior of a Soleil compensator.

Upon comparing the exact phase difference Δ_e with the approximate phase difference Δ_a we found that, for both the calcite Berek and the quartz Ehringhaus rotary compensators, the envelope function $\pm \Phi \sin \Delta_a$ provided a relatively good estimate of the discrepance between Δ_e and Δ_a. Because $|\epsilon - \omega|$ is much larger for calcite than for quartz and because the physical path lengths are an order of magnitude smaller in the Berek compensator compared to the Ehringhaus compensator, the resonance effects prescribe more stringent conditions on beam divergence, monochromaticity and surface flatness for the Ehringhaus compensator than for the Berek compensator. Also, a given phase difference Δ_e can be obtained with many different rotational settings with the Ehringhaus compnesator, while this effect is nearly non-existent in the Berek compensator. For the calcite Berek compensator with $d = 120\lambda_o$, we have that $\partial \Delta_e / \partial i$ is negative over only

[21] The geometrical optics theory of the Soleil compensator is treated in reference 17, p. 691.

one small region of i , for $0^\circ \leqq \Delta_e \leqq 180^\circ$.

An isotropic rotary retardation plate was proposed. A high refractive index material is desired. The proposed device, for a given thickness, has a higher sensitivity for relatively long (infrared) radiation.

A referee has suggested that, "To eliminate internal reflections and therefore the associated resonance effects, all that must be done is to immerse the compensator in an index matching fluid." This is a creative thought and deserves consideration by users and manufacturers of rotary compensators. Achievement of fluid immersion might, of course, require a certain amount of mechanical ingenuity, particularly in instances where a present day rotary compensator is used in an existing commercial polarizing microscope. Because rotary compensators can, in principle, be used at any wavelength for which absorption is negligible, they have potential (as yet, relatively unused) in variable wavelength ellipsometers. Since many ellipsometers are of the "home-made" variety, fluid immersion should be relatively easy to implement. I would like to point out that, for very precise measurements, resonance effects may not be completely negligible, even with fluid immersion, because the anisotropy of the crystal used in the compensator does not permit perfect matching of the p and s components of the incident wave simultaneously.[22]

7. ACKNOWLEDGEMENTS

I wish to thank Drs. R. L. Longini and D. L. Feucht, Carnegie Institute of Technology for support, encouragement, and many helpful suggestions. Merrs. J. Valerio and A. Smith and Mrs. E. Goodworth, assisted in the preparation of this paper. I thank the referees for constructive criticism.

---

[22] Reference 17, pp. 699-700.

FIGURE CAPTIONS FOR APPENDIX C

Fig. C-1 Geometry considered in this work. The xyz coordinate system is assumed to be oriented such that the xy plane is located at the interface of plate 1 and medium I. Media I and II are considered to have unity of refraction. Crystalline plates 1 and 2 are both assumed to have their respective principal axes aligned with the xyz coordinate axes. The input plane wave is incident in the xz plane, with the incident wave normal making an angle i with the z axis. The permeabilities of all media are assumed to be equal to the permeability μ_o of vacuum. Further, we assume that all media are lossless, non optically active, and homogeneous.

Fig. C-2a,b The error angle $\Delta_e - \Delta_a$ in degrees (vertical scale) as a function of the angle of incidence i in degrees (upper horizontal scale) for a calcite Berek compensator. The values of Δ_a are given in degrees (lower horizontal scale) for 2° increments of i. $\Delta_e - \Delta_a$ is plotted for $d = 120\lambda_o$ (——) and $d = 50\lambda_o$ (— — —). The envelope function is also shown (— – —). The wavelength used is 8010Å, for which we have $\omega = 1.64869$ and $\epsilon = 1.48216$.

Fig. C-3a,b The amplitude ratio factor $T = \tan\alpha_o / \tan\alpha_i$ (vertical scale versus i in degrees (horizontal scale) for a calcite Berek compensator with $d = 120\lambda_o$ (——) and $d = 50\lambda_o$ (— — —).

Fig. C-4 Envelope function $\pm 2 \sin \Delta_a$ in degrees (vertical scale) versus angle of incidence i in degrees (horizontal scale) for a quartz Ehringhaus compensator with $d = 1200\lambda_o$ The range $0^\circ \leq i \leq 75^\circ$ covers about $5\frac{1}{2}$ orders of Δ_a at a wavelength of $\lambda_o = 8325$ Å. We have taken

ϵ = 1.54661 and ω = 1.53773. The intervals delineated by the dotted lines are shown in greater detail in Figs. C-5 and C-6.

Fig. C-5 Detail of Fig. C-4 for the range $5.0^\circ \leqq i \leqq 7.5^\circ$. Δ_e and Δ_a are in degrees (vertical scale) and i is in degrees (horizontal scale). The slowly rising curve is the approximate phase difference Δ_a. The two dashed (— — —) curves correspond to the variation of the envelope function, $\pm \Phi \sin \Delta_a$

Fig. C-6 Detail of Fig. C-4 for the range $47.4^\circ \leqq i \leqq 47.6^\circ$. $\Delta_e - \Delta_a$ is in degrees (vertical scale) and i is in degrees (upper horizontal scale). The envelope function $\pm \Phi \sin \Delta_a$ is virtually a straight line. $\Delta_e - \Delta_a$ is shown for $d = 1200.0\,\lambda_o$ (————) and $d = 1200.1\,\lambda_o$ (— — —). The values of the approximate phase difference Δ_a are given in degrees for $d = 1200.0\,\lambda_o$ (middle horizontal scale) and for $d = 1200.1\,\lambda_o$ (lower horizontal scale).

Fig. C-7 Δ_{max} in degrees (vertical scale) versus angle of incidence in degrees (horizontal scale). 1: n = 1.25, 2: n = 1.50, 3: n = 1.75, 4: n = 2.25, 5: n = 3.00, 6: n = 3.75.

Fig. C-8 Wave optics value of phase shift Δ in degrees (vertical scale) versus angle of incidence i in degrees (horizontal scale) for transmission through a thallium bromide iodide optical crystal. The index of refraction at $\lambda_o = 10\mu$ is taken as n = 2.37274.

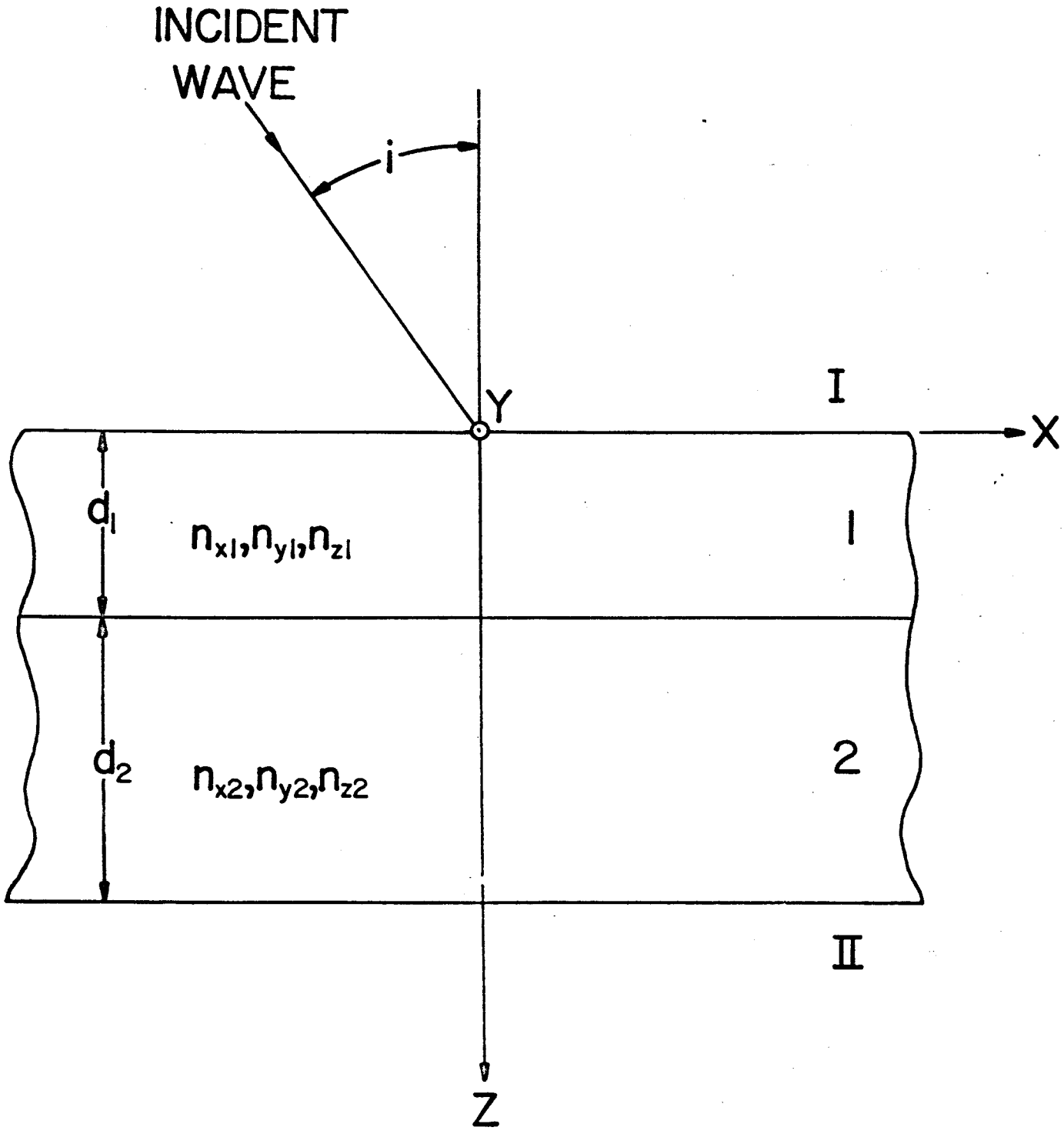


Fig. C-1

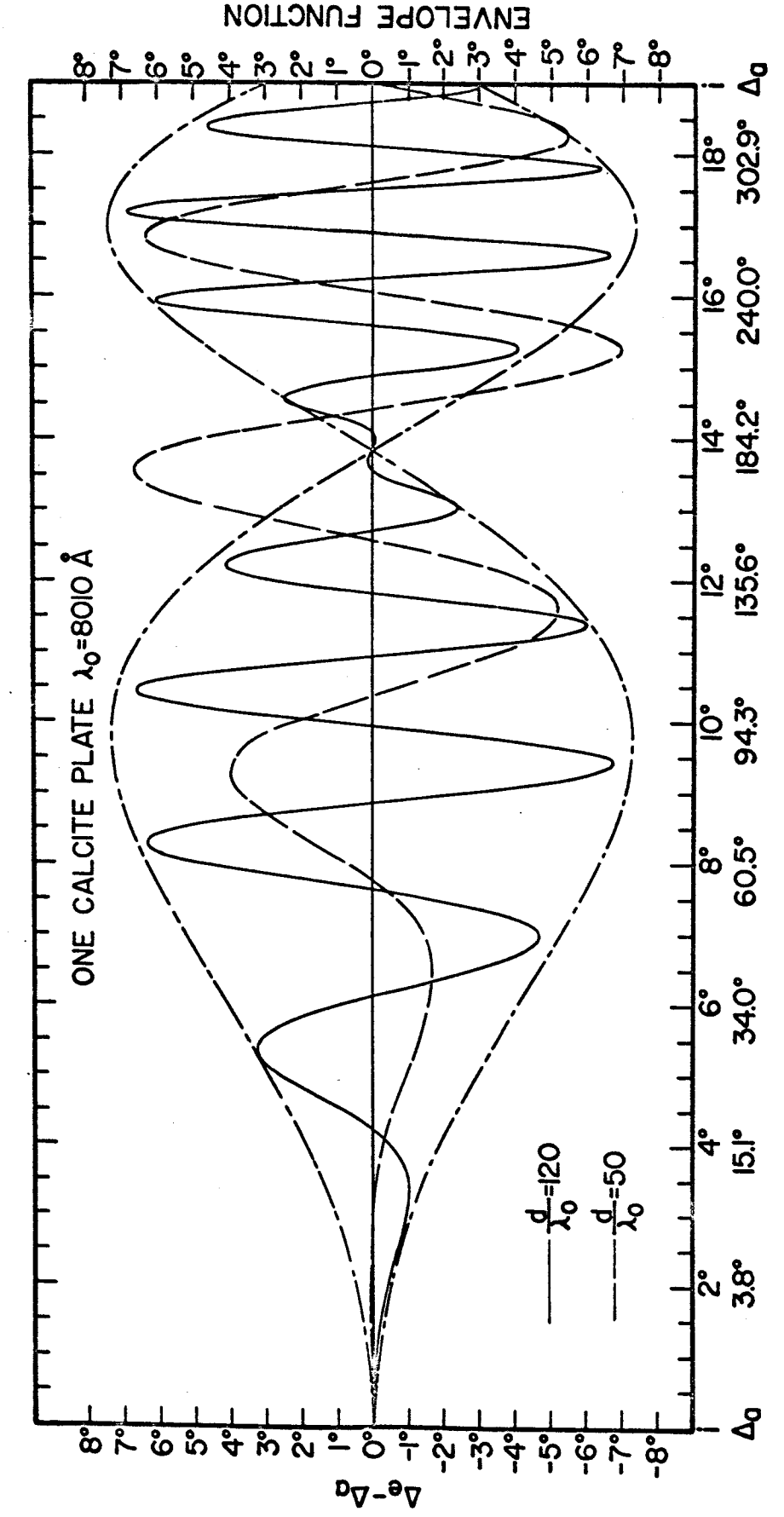

ENVELOPE FUNCTION
ONE CALCITE PLATE λ0=8010 Å
Δe-Δo
Δo
Δe
d/λ0=120
d/λ0=50
8° 7° 6° 5° 4° 3° 2° 1° 0° -1° -2° -3° -4° -5° -6° -7° -8°
2° 3.8°
4° 15.1°
6° 34.0°
8° 60.5°
10° 94.3°
12° 135.6°
14° 184.2°
16° 240.0°
18° 302.9°

Fig. C-2a

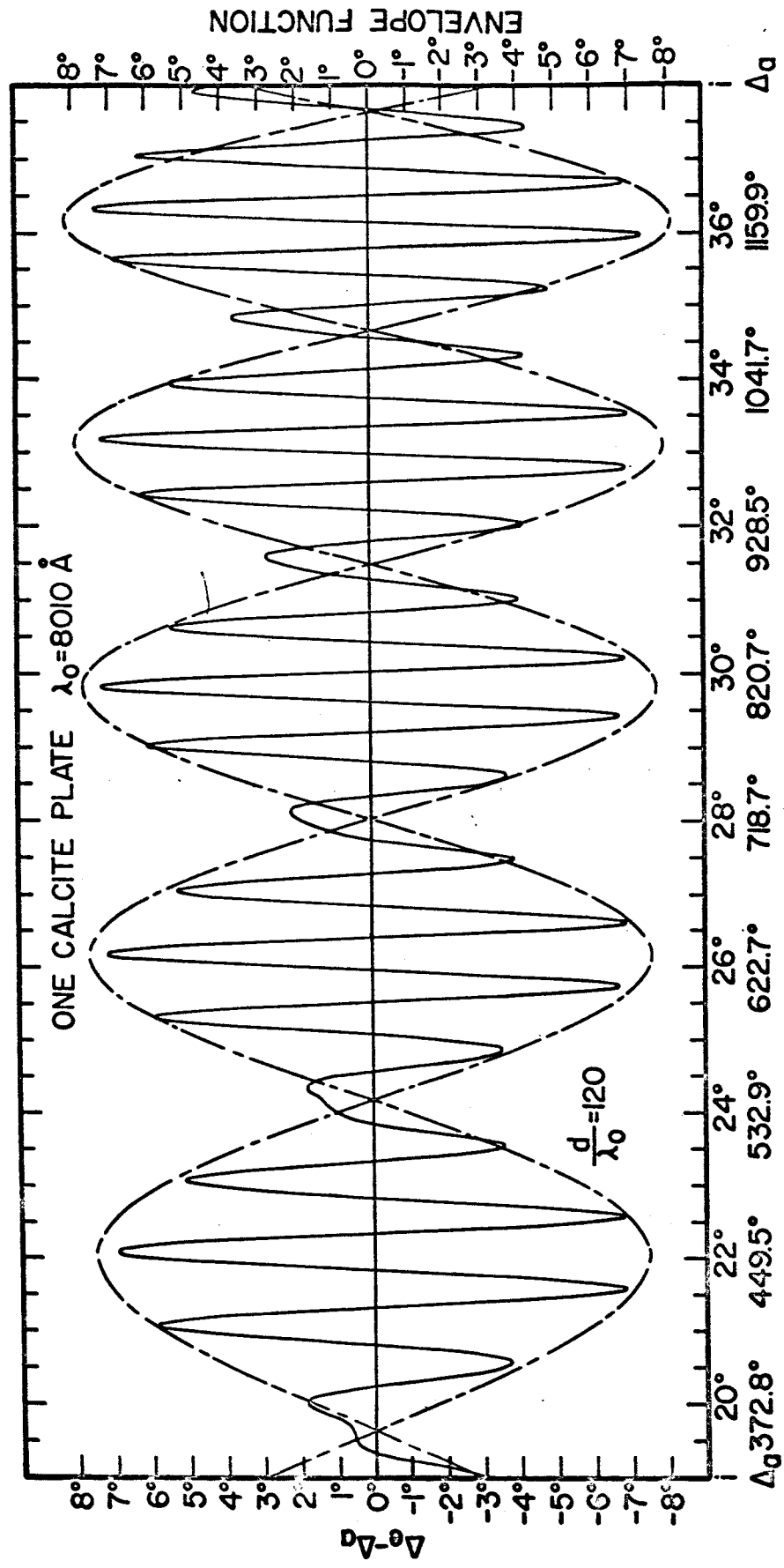

ENVELOPE FUNCTION
ONE CALCITE PLATE λ0=8010 Å
d/λ0=120
Δe-Δo
Δo
372.8°
449.5°
532.9°
622.7°
718.7°
820.7°
928.5°
1041.7°
1159.9°
20°
22°
24°
26°
28°
30°
32°
34°
36°

Fig. C-2b

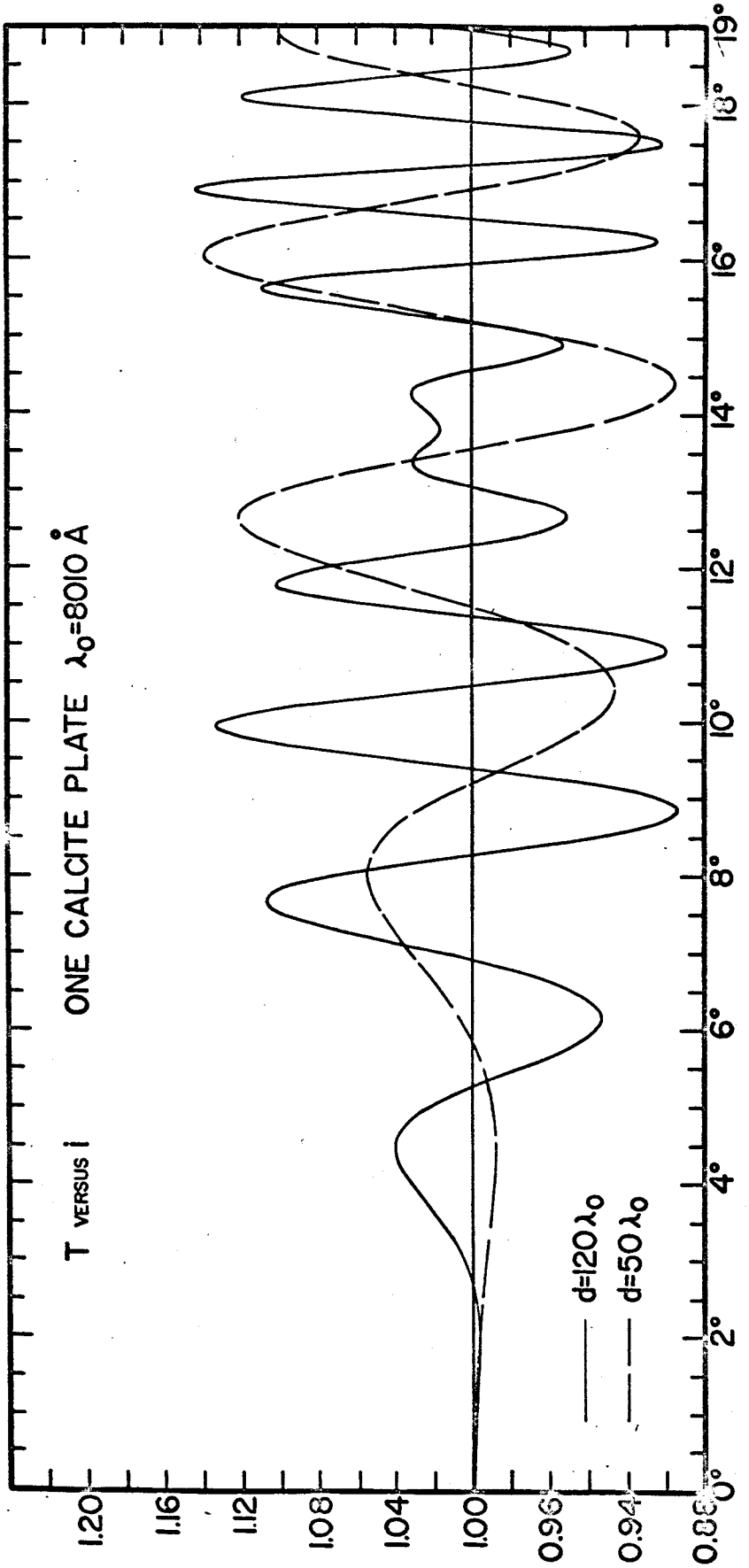

T VERSUS i
ONE CALCITE PLATE λ0=8010 Å
d=120λ0
d=50λ0
1.20
1.16
1.12
1.08
1.04
1.00
0.96
0.94
0°
2°
4°
6°
8°
10°
12°
14°
16°
18°
19°

Fig. C-3a

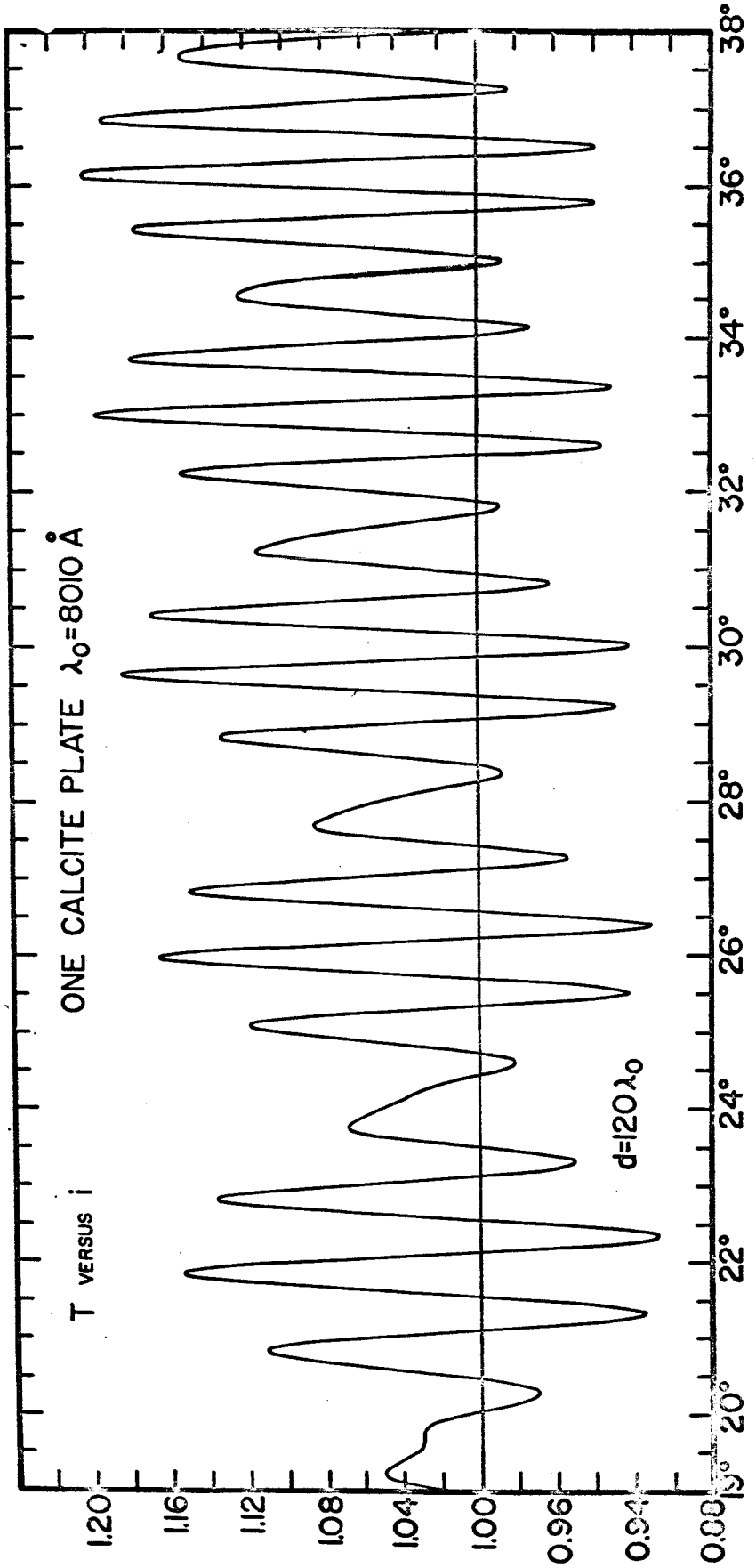

T VERSUS i
ONE CALCITE PLATE λ_0=8010 Å
d=120λ_0
1.20
1.16
1.12
1.08
1.04
1.00
0.96
0.92
0.88
20°
22°
24°
26°
28°
30°
32°
34°
36°
38°

Fig. C-3b

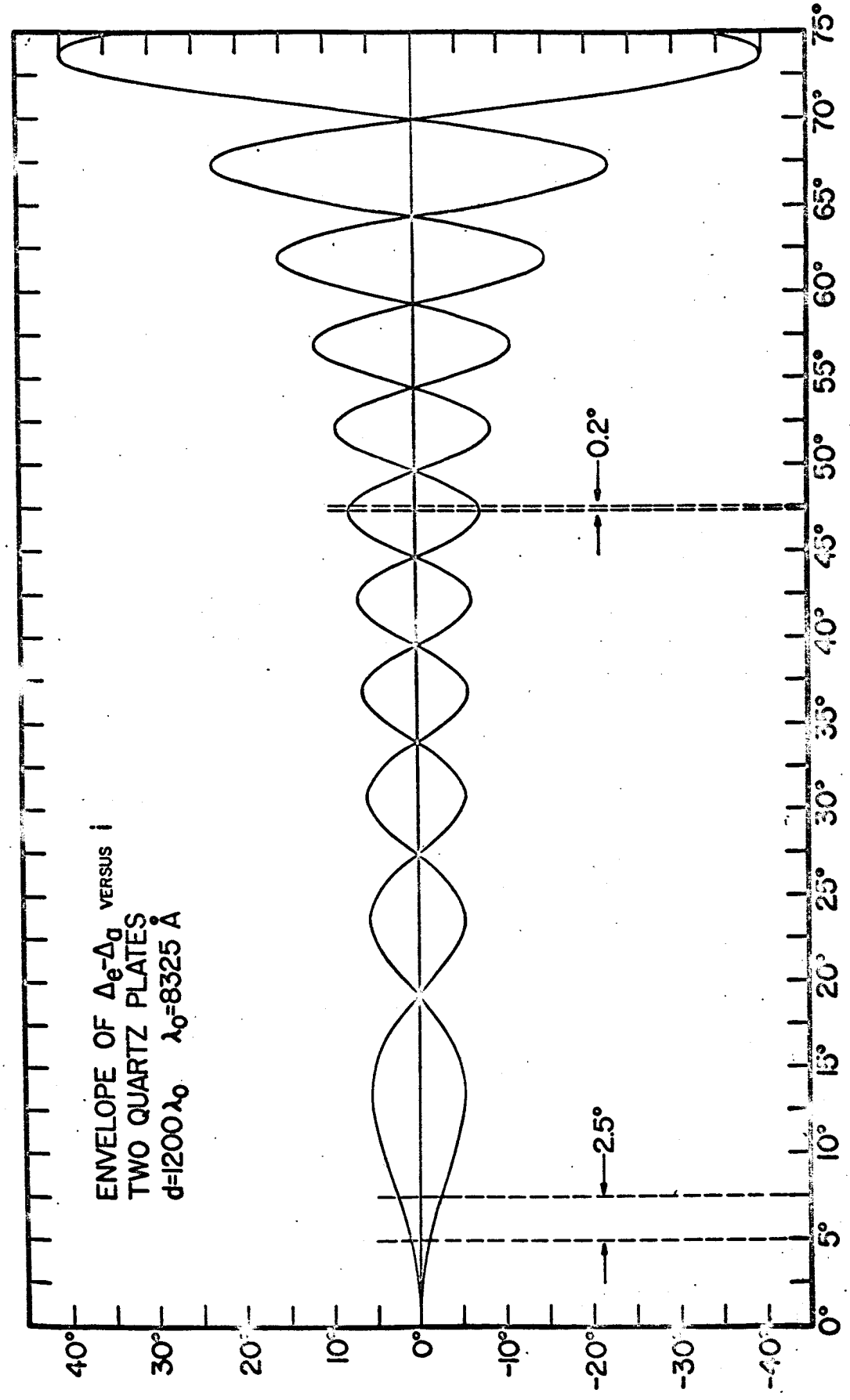

ENVELOPE OF $\Delta_e - \Delta_o$ VERSUS i
TWO QUARTZ PLATES
d=1200λ_o λ_o=8325 Å
0.2°
2.5°
0°
5°
10°
15°
20°
25°
30°
35°
40°
45°
50°
55°
60°
65°
70°
75°
40°
30°
20°
10°
0°
-10°
-20°
-30°
-40°

Fig. C-4

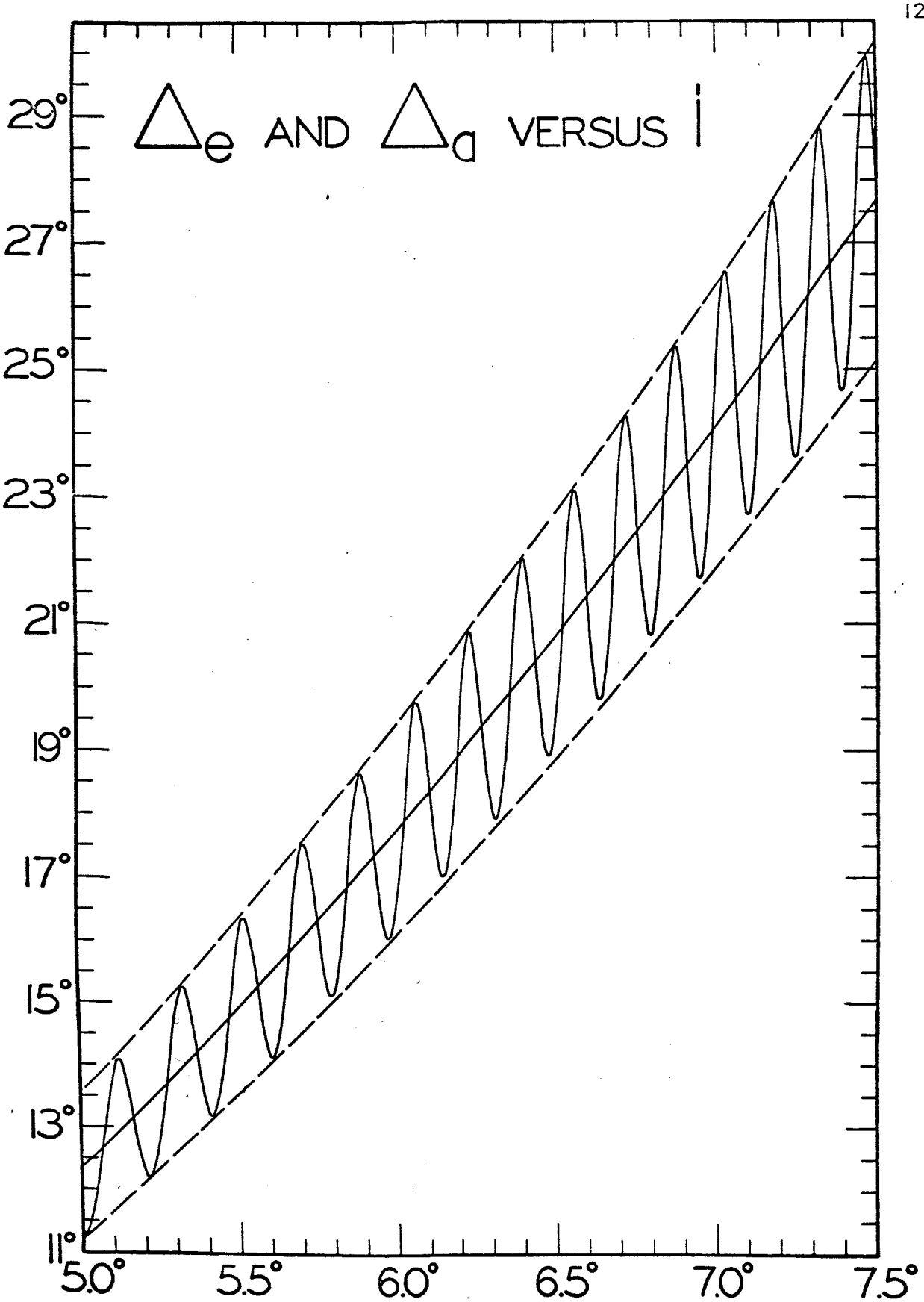

Δ_e AND Δ_a VERSUS i
29°
27°
25°
23°
21°
19°
17°
15°
13°
11°
5.0°
5.5°
6.0°
6.5°
7.0°
7.5°

Fig. C-5

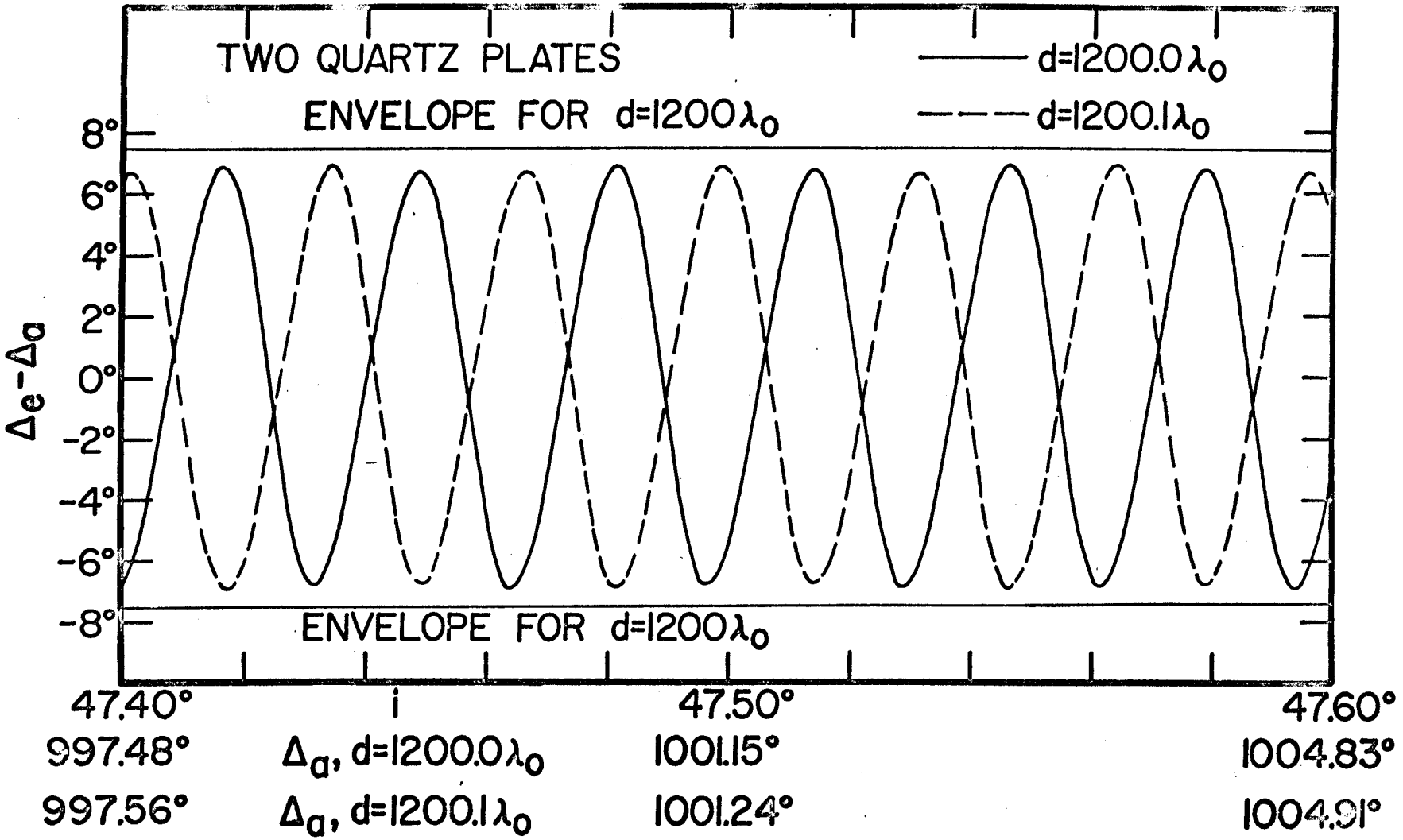

TWO QUARTZ PLATES
d=1200.0λ0
ENVELOPE FOR d=1200λ0
d=1200.1λ0
8°
6°
4°
2°
0°
-2°
-4°
-6°
-8°
Δe-Δa
ENVELOPE FOR d=1200λ0
47.40°
i
47.50°
47.60°
997.48°
Δa, d=1200.0λ0
1001.15°
1004.83°
997.56°
Δa, d=1200.1λ0
1001.24°
1004.91°

Fig. C-6

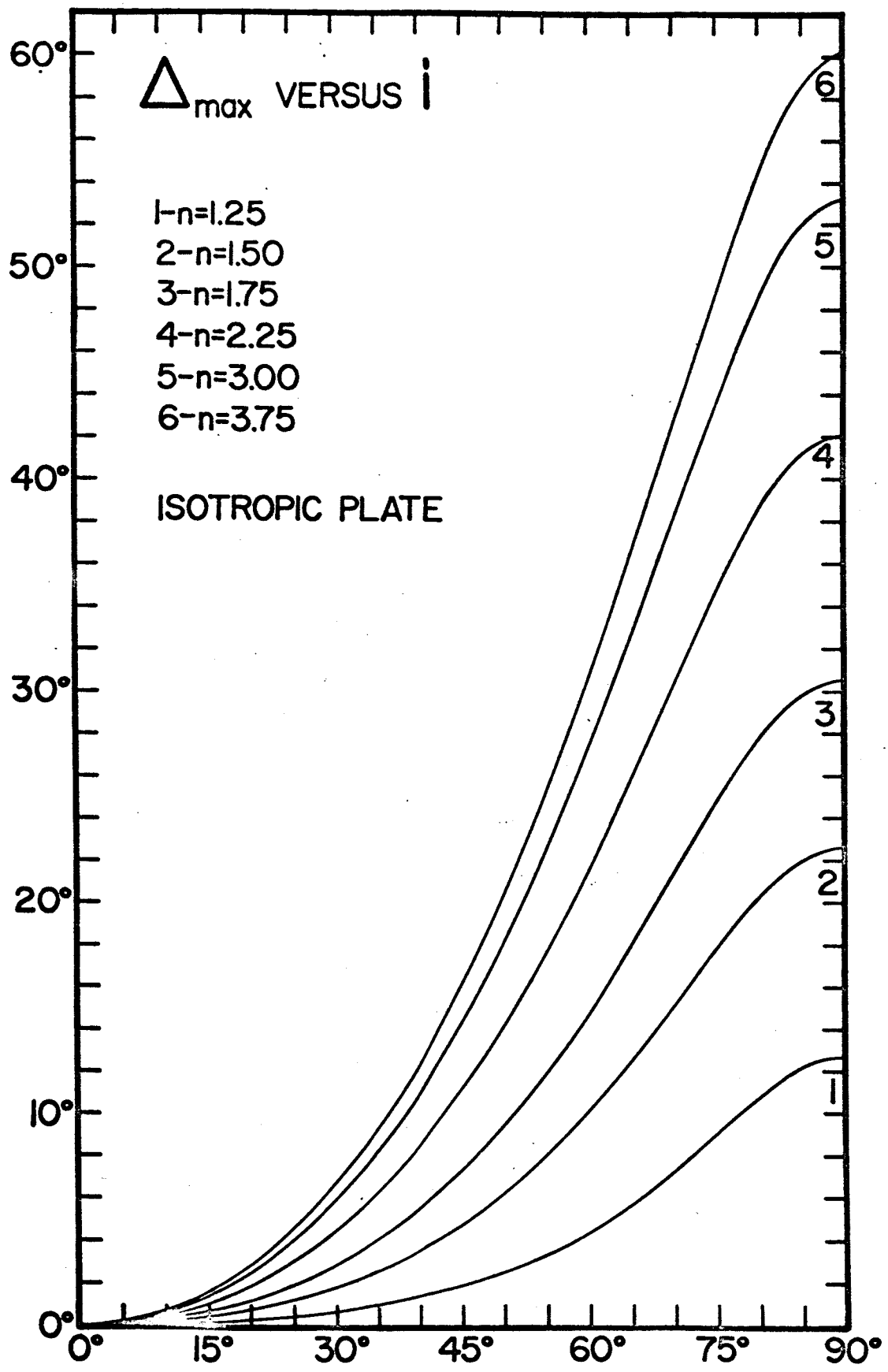

Δmax VERSUS i
1-n=1.25
2-n=1.50
3-n=1.75
4-n=2.25
5-n=3.00
6-n=3.75
ISOTROPIC PLATE
60°
50°
40°
30°
20°
10°
0°
0°
15°
30°
45°
60°
75°
90°
1
2
3
4
5
6

Fig. C-7

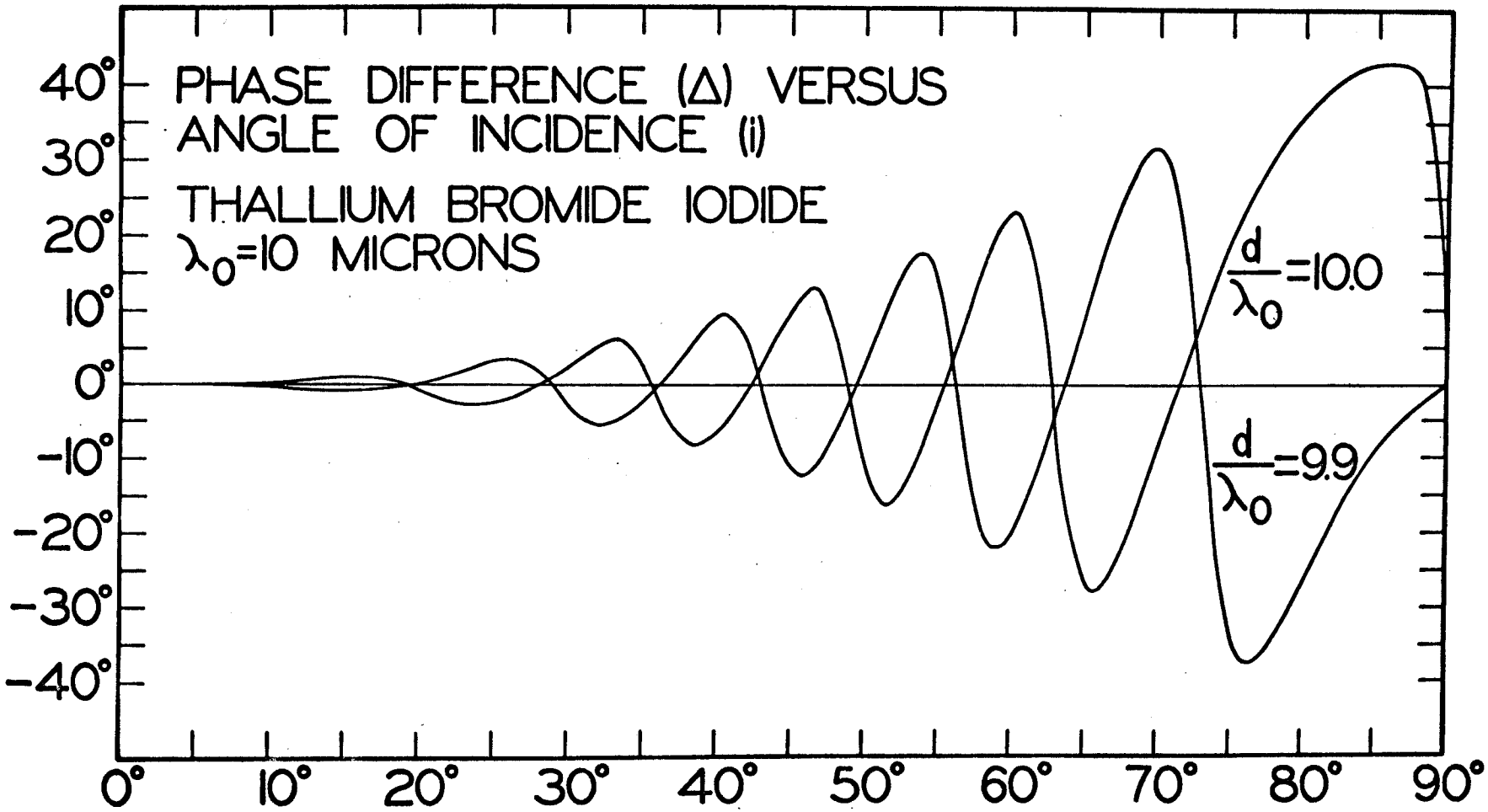

PHASE DIFFERENCE (Δ) VERSUS
ANGLE OF INCIDENCE (i)
THALLIUM BROMIDE IODIDE
λ0=10 MICRONS
d/λ0=10.0
d/λ0=9.9
40°
30°
20°
10°
0°
-10°
-20°
-30°
-40°
0°
10°
20°
30°
40°
50°
60°
70°
80°
90°

Fig. C-8

APPENDIX D

On Fluid Immersion as a Means for Reducing Resonance Effects is Rotary Compensators

This appendix, which is a continuation of App. C, has been published under a similar title in the Journal of the Optical Society of America, vol. 55, p. 209 (1965). The differences between this appendix and the published article are minor. Permission to reproduce the published work has been granted by the Journal of the Optical Society of America. The figures and figure captions relevant to this appendix can be found in numerical sequence at the end of the appendix.

On Fluid Immersion as a Means for Reducing Resonance Effects in Rotary Compensators*

D. A. Holmes

Department of Electrical Engineering
Carnegie Institute of Technology
Pittsburgh, Pennsylvania 15213

In a recent paper[1], some of the electromagnetic aspects of rotary compensators were discussed. There the compensator was considered to be operated in an air environment; the present letter will highlight some of the effects to be expected when the compensator is immersed in an index matching fluid[2], using the calcite Berek rotary compensator for illustrative purposes. While we shall consider only the calcite Berek compensator in this letter, the use of an index matching environment offers possibilities for reducing undesirable multiple reflection effects in wave-plates[3,4] and other types of optical compensators.

When the calcite Berek compensator is surrounded by a medium with refractive index n, the wave optics formulae for the amplitude ratio factor T and the phase difference Δ_e are still given by eqtns. (22) and (23) of the previous work[1] provided that the following substitutions are made:

*This work was supported by the National Aeronautics and Space Administration under contract NAS8-5269 and is contained in a dissertation presented by the author to Carnegie Institute of Technology in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

---

[1] D. A. Holmes, J. Opt. Soc. Am. 54, 1340 (1964), (App. C).

[2] Fluid immersion as a means of eliminating resonance effects in rotary compensators was suggested by a referee of the previously cited work.[1]

[3] H. Weinberger and J. Harris, J. Opt. Soc. Am. 54, 552 (1964).

[4] D. A. Holmes, J. Opt. Soc. Am. 54, 1115 (1964), (App. B).

$$\beta_s = (2\pi/\lambda_0)(\omega^2 - n^2\sin^2 i)^{1/2}\ ,\ \beta_p = (2\pi\omega/\epsilon\lambda_0)(\epsilon^2 - n^2\sin^2 i)^{1/2},$$

$$2K_p = \frac{(\epsilon\omega/n)\cos i}{(\epsilon^2 - n^2\sin^2 i)^{1/2}} + \frac{(\epsilon^2 - n^2\sin^2 i)^{1/2}}{(\epsilon\omega/n)\cos i}\ , \tag{1}$$

$$2K_s = \frac{n\cos i}{(\omega^2 - n^2\sin^2 i)^{1/2}} + \frac{(\omega^2 - n^2\sin^2 i)^{1/2}}{n\cos i}\ .$$

The geometrical optics value of the phase difference is $\Delta_a = (\beta_s - \beta_p)d$.

As we saw in the previous work, the envelope function $\pm\Phi\sin\Delta_a$ provided a good estimate of the gross variation of $\Delta_e - \Delta_a$. It is clear that, if n can be adjusted to a value such that $\Phi = (K_p + K_s - 2)/2$ is smaller than the value of Φ for $n = 1$, then the value of $\Delta_e - \Delta_a$ will be correspondingly smaller for the fluid immersion case than for air immersion. K_p and K_s can alternatively be written as

$$K_p = (1 + r_p^2)/(1 - r_p^2)\ ,\quad K_s = (1 + r_s^2)/(1 - r_s^2)\ , \tag{2}$$

where

$$r_s = \frac{n\cos i - (\omega^2 - n^2\sin^2 i)^{1/2}}{n\cos i + (\omega^2 - n^2\sin^2 i)^{1/2}}\ , \tag{3}$$

$$r_p = \frac{\epsilon\omega\cos i - n(\epsilon^2 - n^2\sin^2 i)^{1/2}}{\epsilon\omega\cos i + n(\epsilon^2 - n^2\sin^2 i)^{1/2}}\ . \tag{4}$$

The condition on n for Φ to be small can thus be translated into the physically reasonable requirement that n be adjusted so that the Fresnel coefficients r_s and r_p both become small. Because the optic axis of the calcite plate is perpendicular to the plate surfaces, the s component of the incident wave undergoes ordinary refraction at all angles of incidence. By choosing the index of the fluid as $n = \omega$, the plate becomes invisible to the s component, i.e., $r_s = 0$, and

multiple reflections are thereby eliminated. The p component of the incident electric field vector, however, is extraordinarily refracted by the calcite plate, that is, the index of refraction "seen" by the p component is a function of the angle of incidence, hence, no value of n will make the plate invisible to the p component for all angles of incidence.

Although $\Delta_e - \Delta_a$ cannot be made identically zero for all angles of incidence it is possible to choose values for n that will make the differences between Δ_e and Δ_a negligible in most experimental situations. Fig. D-1 shows Φ versus the angle of incidence i for some selected values of n . For curves 4, 5, and 6, $|\Delta_e - \Delta_a|$, which is smaller than Φ , is practically zero. The range considered in the angle of incidence corresponds to a range in Δ_a , for the case $n=1$, of $0^\circ \leqq \Delta_a \leqq 1{,}100^\circ$.

While the deleterious resonance effects can be virtually eliminated by judicious choice of an index matching fluid we should recognize that fluid immersion necessitates recalibration of the compensator. Recalibration is necessary because the phase difference Δ_a is a function of the angles of refraction (of the respective p and s components) which, in turn, depend upon the refractive index of the surrounding medium. The usual calibration charts supplied with a rotary compensator are based on an air environment. Also fluid immersion decreases the precision of the compensator, that is to say, if δi is the rotational setting accuracy, then the uncertainty in Δ_a , which is given by $(\partial \Delta_a / \partial i) \cdot \delta i$, becomes larger. These features are illustrated in Fig. D-2. Finally, we should recognize that the refractive index of an immersion medium can be temperature sensitive, hence, temperature control of the immersion medium may prove desirable in an experimental situation. By referring to Fig. D-2, we note that if the index of the surrounding media changes from an initial value of $n = 1.64869$ to $n = 1.7$ then the corresponding value of Δ_a shifts from $\Delta_a = 93^\circ$ to $\Delta_a = 99^\circ$, for $i = 6^\circ$. For operation at high angles of incidence, the stability of the index of the surrounding medium assumes greater importance.

I wish to thank professors R. L. Longini and D. L. Feucht, Carnegie Institute of Technology, for several beneficial discussions.

FIGURE CAPTIONS FOR APPENDIX D

Fig. D-1 Theoretical variation of $\Phi = (K_p + K_s - 2)/2$ (vertical scale, degrees) with angle of incidence i (horizontal scale, degrees). Recall that $\Delta_e \cong \Delta_a + \Phi \cos\theta \sin\Delta_a$, hence, $|\Delta_e - \Delta_a|$ is generally much smaller than Φ. The numbering of the curves specifies the index of refraction of the fluid as follows: 1: n = 1.45; 2: $n = \epsilon = 1.48216$; 3: n = 1.55; 4: n = 1.60; 5: $n = \omega = 1.64869$; 6: n = 1.70. The parameters for the calcite plate were taken as $d = 120\lambda_o$, ϵ = 1.48216 (extraordinary refractive index) and ω = 1.64869 (ordinary refractive index).

Fig. D-2 Theoretical variation of the geometrical optics value of phase difference Δ_a (degrees) versus angle of incidence i (degrees). The index of the surrounding medium is specified by the curve numbers similarly to Fig. D-1. For comparison, curve 0 refers to $n = 1$, or an air environment. The parameters for the calcite plate are the same as for Fig. D-1.

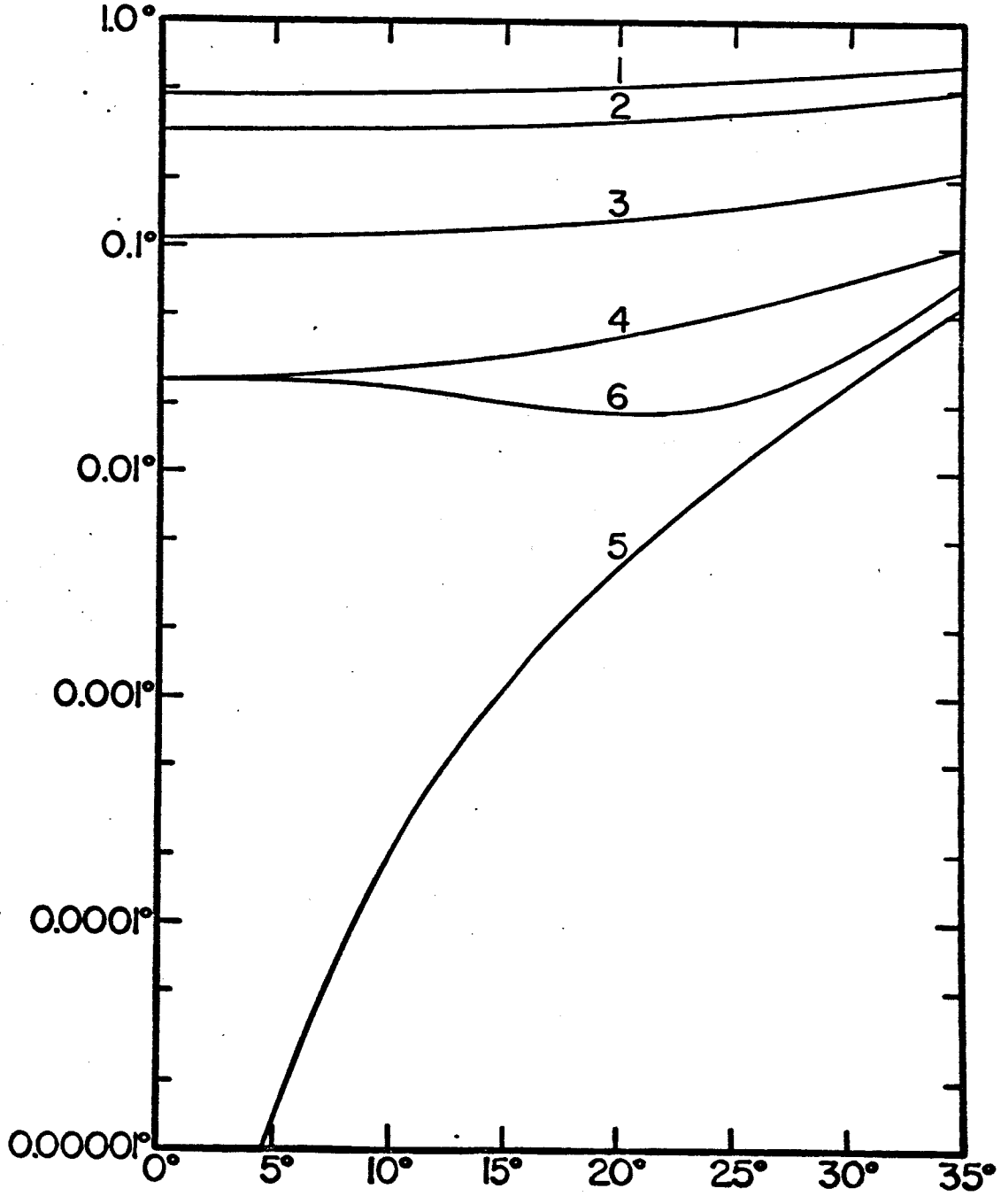


Fig. D-1

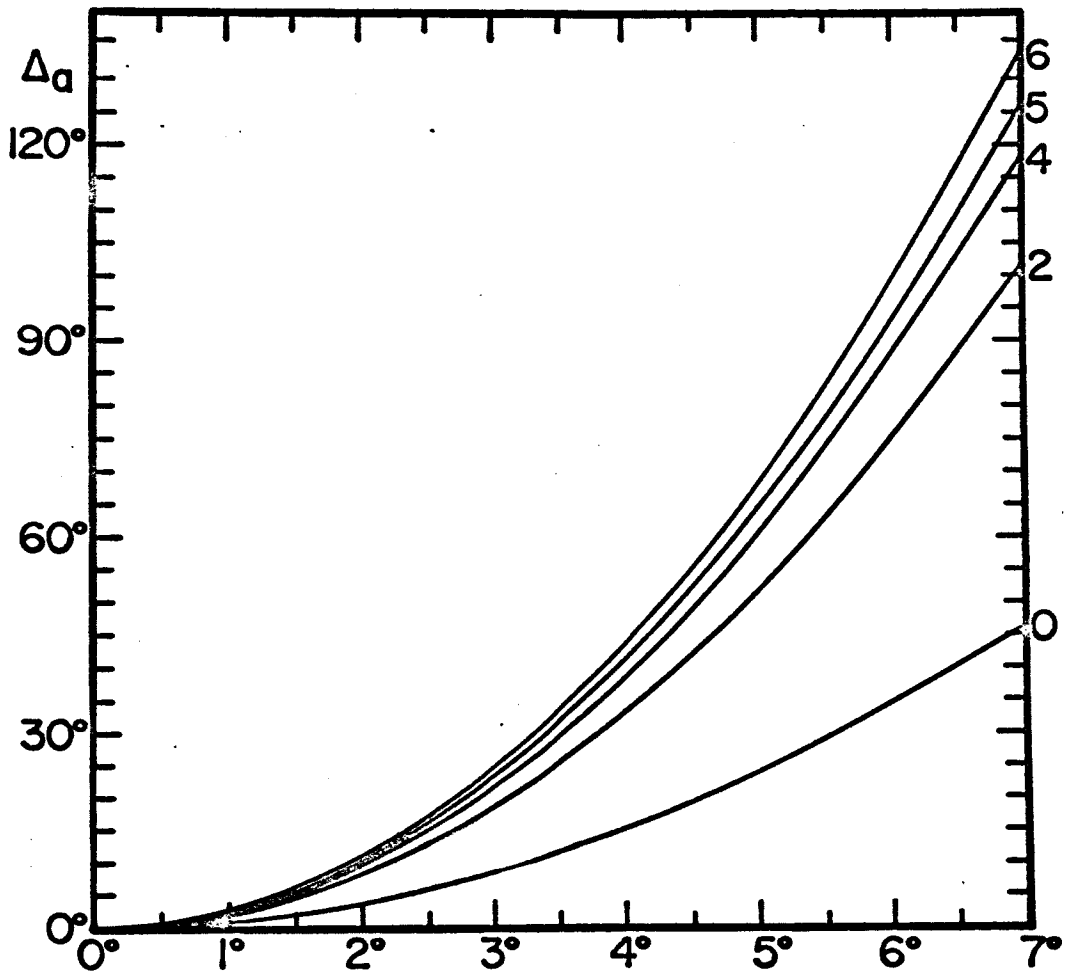


Fig. D-2

APPENDIX E

Error Analysis for Ellipsometric Measuring Techniques

In any experiment, measurement errors exist which, in turn, influence the accuracy of calculated quantities. In this appendix I consider measurement errors in the ellipsometric investigation of weakly absorbing isotropic substances. Of particular interest is the manner in which a measurement error propagates in the inversion relations (26) and (32a,b,c).

E.1. TWO MEDIA PROBLEM (ISOTROPIC GaAs)

In the measurement of i , Δ , and ψ one has to expect a certain amount of experimental error. We shall let i , Δ , and ψ stand for true values and δi , $\delta\Delta$, and $\delta\psi$ stand for the respective measurement errors. The error $\delta\psi$ actually represents the calculated error in ψ when $\psi_i + \delta\psi_i$ and $\psi_r + \delta\psi_r$ are substituted into (23). In the following paragraphs we consider how the measurement errors influence the accuracy in calculating the optical constants. We shall not consider multiple errors; instead, we shall consider a variation in only one measurement at a time while the remaining measured quantities are held at their true values. Our approach shall be in the form of numerical examples rather than be a general analytical error analysis. We might remark, however, that inspection of (26) indicates that n and k could be expanded to first order in δi , $\delta\Delta$, and $\delta\psi$ by using such approximations as $\sin(i+\delta i) \cong \sin i + \delta i \cos i$. We shall begin by assuming various true values of i , n , and k and then calculate the true values of Δ and ψ . Then, after assuming numerical values for the measurement errors δi , $\delta\Delta$, and $\delta\psi$, the calculated values, n_c and k_c , of the optical constants can be determined by substituting $i+\delta i$, $\Delta+\delta\Delta$, and $\psi+\delta\psi$ into the inversion formula (26). We shall consider true values of i which are close to the Brewster angle when the true value $n = 3.58$ with the true value of k in the range $.001 \leqq k \leqq .007$. Further we shall assume that the precision of the measurements* allows

*The maximum measurement errors considered here are felt to be larger than those characterizing our ellipsometer.

the restrictions that $|\delta i| \leqq .06^\circ = 3.6'$, $|\delta\Delta| \leqq 0.45^\circ = 27'$, and $|\delta\psi| \leqq 0.15^\circ = 9'$.

First we consider the error δi, assuming that $\delta\Delta = 0$ and $\delta\psi = 0$. From numerical computations we find that $n_c = n + 0.15(\delta i/3.6)$ and $k_c = k[1 + 4.6\times10^{-6}(\delta i/3.6)]$ where δi is in minutes of arc. These expressions are true for $74.3^\circ \leqq i \leqq 74.5^\circ$, $n = 3.58$, and $.001 \leqq k \leqq .007$. We should expect n_c and k_c to be relatively independent of the precise value of i within the range indicated because over this particular and small range of angles the trigonometric functions of i are practically constant.

Second we consider the error $\delta\Delta$, assuming that $\delta i = 0$ and $\delta\psi = 0$. For a transparent medium, i.e., $k = 0$, $\Delta = \pi$ for $i < i_B$ and $\Delta = 0$ for $i > i_B$, hence, any deviation from $\Delta = \pi$ or $\Delta = 0$ occurs solely because $k > 0$. The exact value of Δ depends, of course, upon n, i, and k and, as we saw earlier, Δ is a sensitive function of k when i is close to $\tan^{-1}(n)$. Fig. E-1 shows how k_c depends upon $\delta\Delta$. Curves 1 correspond to $i_1 = 74.3^\circ$ while curves 2 correspond to $i_2 = 74.4^\circ$. The Brewster angle is $i_B = 74.39338^\circ$, hence i_2 is closer to i_B than i_1 and, from the figure, it is apparent that a given $\delta\Delta$ produces a greater error in k_c when the measurement is taken at i_1 than when the measurement is taken at i_2. The error $\delta\Delta$ is almost wholly absorbed by k_c; the calculations reveal that, for both angles of incidence and both values of k, n_c differs from n in the sixth significant figure. The reason for this is that k_c is not (percentage-wise) much different from k, thus, only a minute change in n_c from n is required in (25) to yield the true value of ψ. From a purely mathematical standpoint, we see from (26) that, when ψ is very small and when $|\delta\Delta| << |\Delta|$, the effect of $\delta\Delta$ on n_c is negligible.

Third we consider the error $\delta\psi$, assuming $\delta i = 0$ and $\delta\Delta = 0$, using the same values for i, n, and k that were used in the previous paragraph. Fig. 3-3 shows that ψ is very small, e.g., curve 3 gives a value of $\psi \cong 11'$, hence, an error of $\delta\psi = 9'$ results in a situation where the experimental error $\delta\psi$ is nearly as large as the true value of ψ. For curve 5 in Fig. 3-3, $\delta\psi = 9'$ is larger than ψ. This is a direct consequence of choosing i near the Brewster angle, a

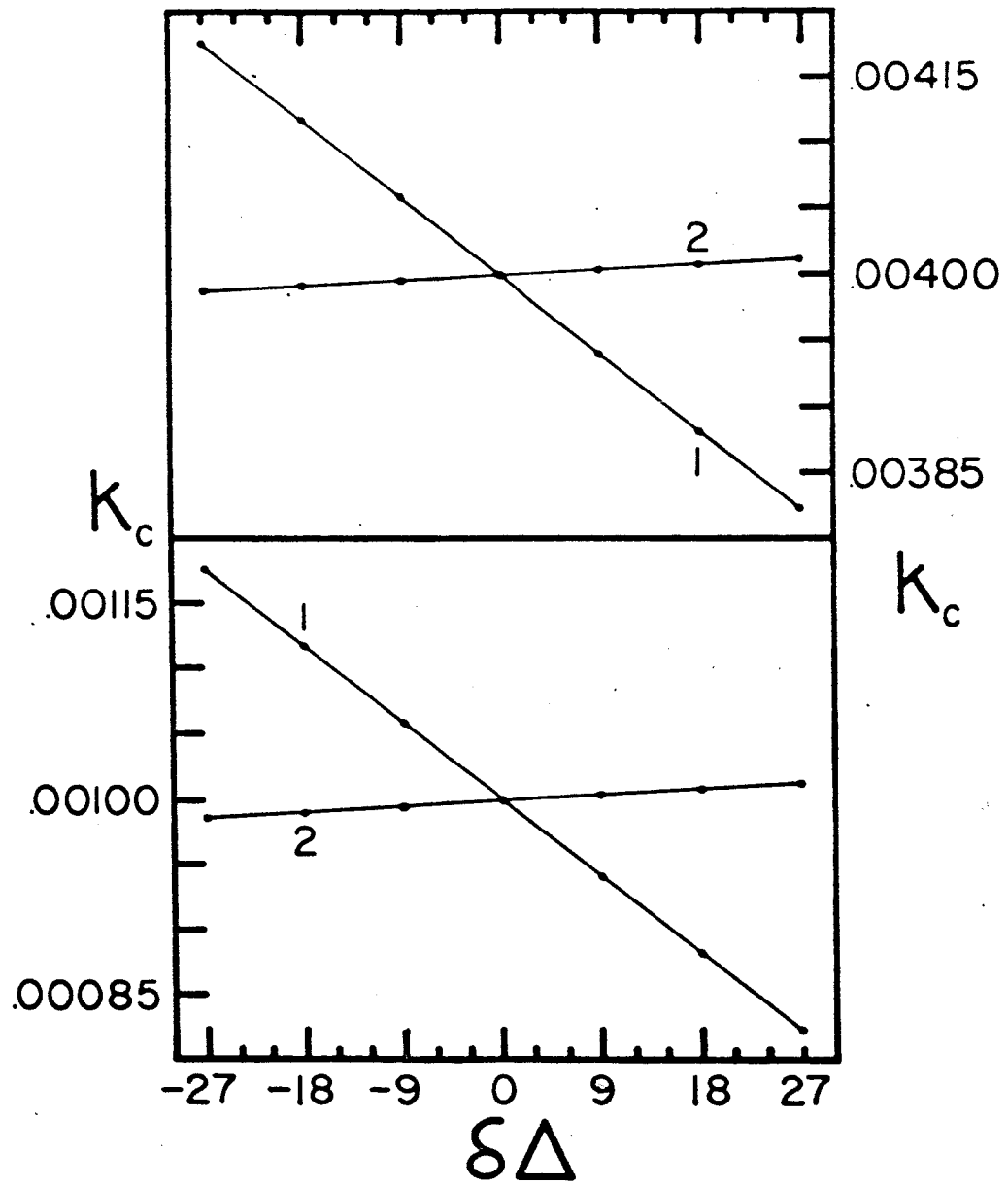


Fig. E-1 Variation of k_c with $\delta\Delta$ for the two media geometry, assuming $\delta i = \delta\psi = 0$. The scale for $\delta\Delta$ is in minutes of arc. Curves 1 correspond to an angle of incidence $i_1 = 74.3^\circ$, and curves 2 correspond to an angle of incidence $i_2 = 74.4^\circ$, while the Brewster angle is $i_B = 74.39338^\circ$. A perfect measurement of Δ , i.e., $\delta\Delta = 0$, yields $k_c = k$. The dots (•) indicate the calculated points.

choice we were forced to make in order to make Δ and ψ sensitive to variations in k. Since $\tan(\psi+\delta\psi)$ is still a very small quantity, even when $\delta\psi = 9'$, n_c cannot wander too far from n, therefore the error $\delta\psi$ has to manifest itself predominantly as an error in k_c. By examining equation (26) we see that the error in k_c is approximately proportional to $4\delta\psi$ and further, when $\psi+\delta\psi<0$, we will have $k_c<0$! The complex number n_c+jk_c will travel along a path in (n_c, k_c) space as $\delta\psi$ is varied, as shown in Fig. E-2. The main point established by Fig. E-2 is that the error in k_c is probably not tolerable for physically realizeable values of $\delta\psi$. By moving i away from the Brewster angle thus increasing ψ, the ratio $|\delta\psi/\psi|$ can be reduced, which correspondingly will reduce the error in k_c. Moving i away from the Brewster angle, however, reduces the sensitivity of Δ to k and the error in k_c due to $\delta\Delta$ becomes larger as shown in Fig. E-1. Also, as i moves away from i_B, either of the quantities $\pi-\Delta$ or Δ can become small, for example, for $i = 74.0^\circ$, $n = 3.58$, and $k = .001$ we have $\pi-\Delta = 0.61^\circ$, hence, i does not have to be far removed from i_B to make $\delta\Delta$ comparable to $\pi-\Delta$ or Δ.

From the above discussions it is clear that a two media investigation of a weakly absorbing substance poses some conflicting alternatives. Operation very close to the Brewster angle causes Δ and ψ to become sensitive to k, and Δ is sufficiently far removed from π or 0 so that $\delta\Delta$ is very small compared to Δ, however, ψ is so small that $\delta\psi$ can easily become comparable to ψ. Moving away from the Brewster angle increases ψ, reducing the effect of $\delta\psi$, however, Δ and ψ become less sensitive to k and, in addition, Δ can approach π or 0, causing $\delta\Delta$ to become comparable to $\pi-\Delta$ or Δ. The enduring question, which we have not really answered, is, what is the optimum value of i in order to minimize the errors in n_c and k_c due to $\delta\Delta$ and $\delta\psi$ (and combinations thereof)? The answer to this question would probably emerge from an analytical error analysis, however, the preceding discussion, including the numerical examples, lends credence to the thought that, even if an optimum value of i is found, the remaining errors in n_c and, particularly, k_c, would still be too large to make the two media approach attractive.

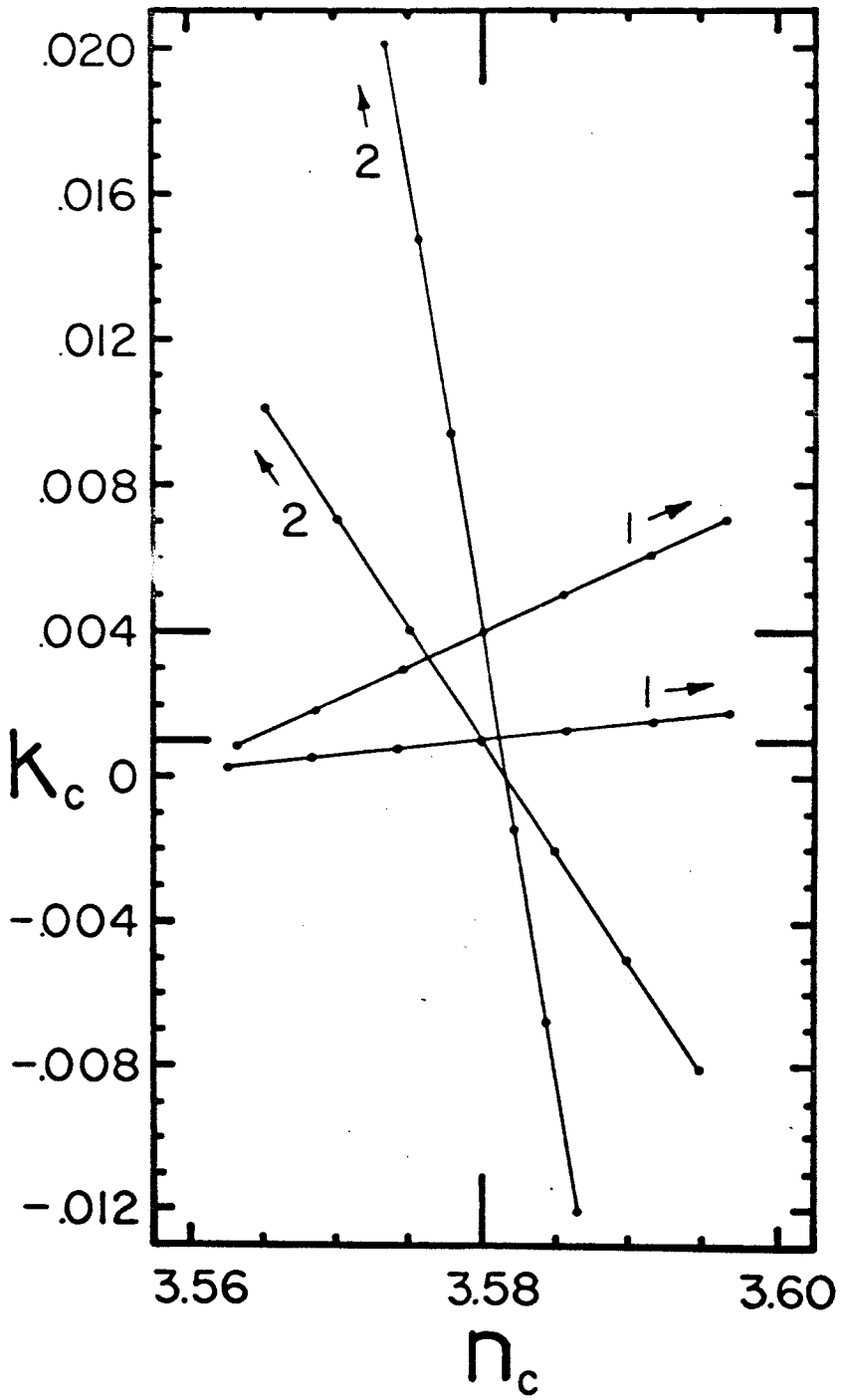


Fig. E-2 A plot of the complex quantity $n_c + jk_c$ in (n_c , k_c) space with $\delta\Psi$ as the changing parameter, assuming $\delta i = \delta\Delta = 0$. Curves 1 are for $i = 74.3^\circ$ and curves 2 are for $i = 74.4^\circ$, while the arrows indicate the direction of increasing $\delta\Psi$. Points were calculated for $\delta\Psi = \pm 9'$, $\pm 6'$, $\pm 3'$, 0 and are indicated by dots (.). When $\delta\Psi = 0$ we have $k_c = k$ and $n_c = n$ for the two media geometry under consideration. Note that two values of k were assumed ($k = 0.001$, $k = 0.004$).

E.2. THREE MEDIA PROBLEM (ISOTROPIC GaAs)

In a numerical example to illustrate the use of (32) we shall suppose (arbitrarily) that measurements have been taken at $i_1 = 54^\circ$ and $i_2 = 56^\circ$. In addition to the mentioned true values for i_1 and i_2 we shall assume that the true values Δ_1 , Δ_2 , Ψ_1 , and Ψ_2 exist for the true values $n = 3.58$, $N = 0.2$, $K = 30$, $d = 50\mu$, $\lambda = 0.84\mu$, with the two true values $k = .001$ and $k = .004$. The maximum errors for the measured quantities shall be assumed to be $|\delta i_1| = |\delta i_2| = 0.06^\circ$, $|\delta\Delta_1| = |\delta\Delta_2| = 0.45^\circ$, $|\delta\Psi_1| = |\delta\Psi_2| = 0.15^\circ$, $\delta N = .06$, and $\delta K = 9.0$. The errors in N and K can thus be as high as 30%. For convenience we set $\delta\lambda = 0$.

The behavior of Δ and Ψ with variations in k is shown in Figs. E-3 and E-4. Note that Δ and Ψ become less sensitive to changes in k when k becomes large. For the largest values of k shown, Δ approaches π , implying that the three media geometry is beginning to look like a two media geometry to the incident wave. From inspection of the figures we should expect the percentage errors in k_c due to $\delta\Delta$ and $\delta\Psi$ to be larger for $k = .004$ than for $k = .001$. As in the two media case, we shall consider only one error at a time assuming that the remaining measurements are free of error.

The calculated quantities $n_c + jk_c$ and d_c are shown in Figs. E-5 and E-6. The errors due to δN and δK are not large enough to be significant on the scales of the figures. This is because the silver substrate with $n_3 = 0.2 - j\,30$ is nearly a perfect reflector and even with 30% errors in N and K the substrate is still a nearly perfect reflector, hence, we can conclude that, for a high reflectivity substrate, the determination of the optical properties of the substrate need not be precise. The remaining measurement errors contribute to errors in n_c , k_c , and d_c in ways which cannot easily be explained because of the complicated nature of the inversion formulae (32a,b,c). The most significant observation that can be made is that the three media approach appears to be more suitable than the two media approach if one is mainly interested in obtaining a good estimate of k only.

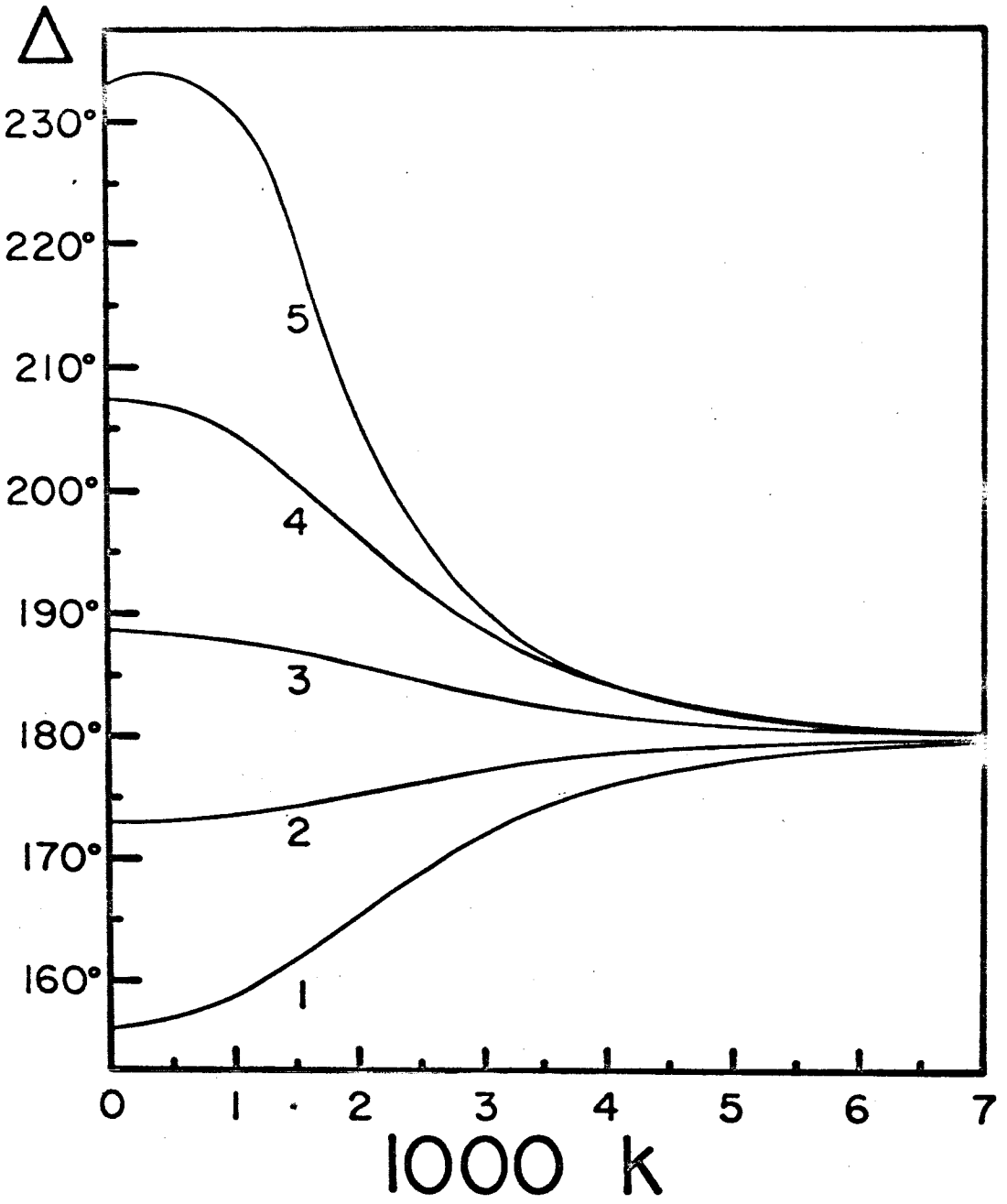


Fig. E-3 Plot of theoretical values of Δ (degrees) versus $1000k$; the three media problem. The numbering of the curves corresponds to the angle of incidence as follows. (1) $i = 54.0^{\circ}$, (2) $i = 54.5^{\circ}$, (3) $i = 55.0^{\circ}$, (4) $i = 55.5^{\circ}$, and (5) $i = 56.0^{\circ}$. The remaining parameters have the same values as for Fig. 3-4.

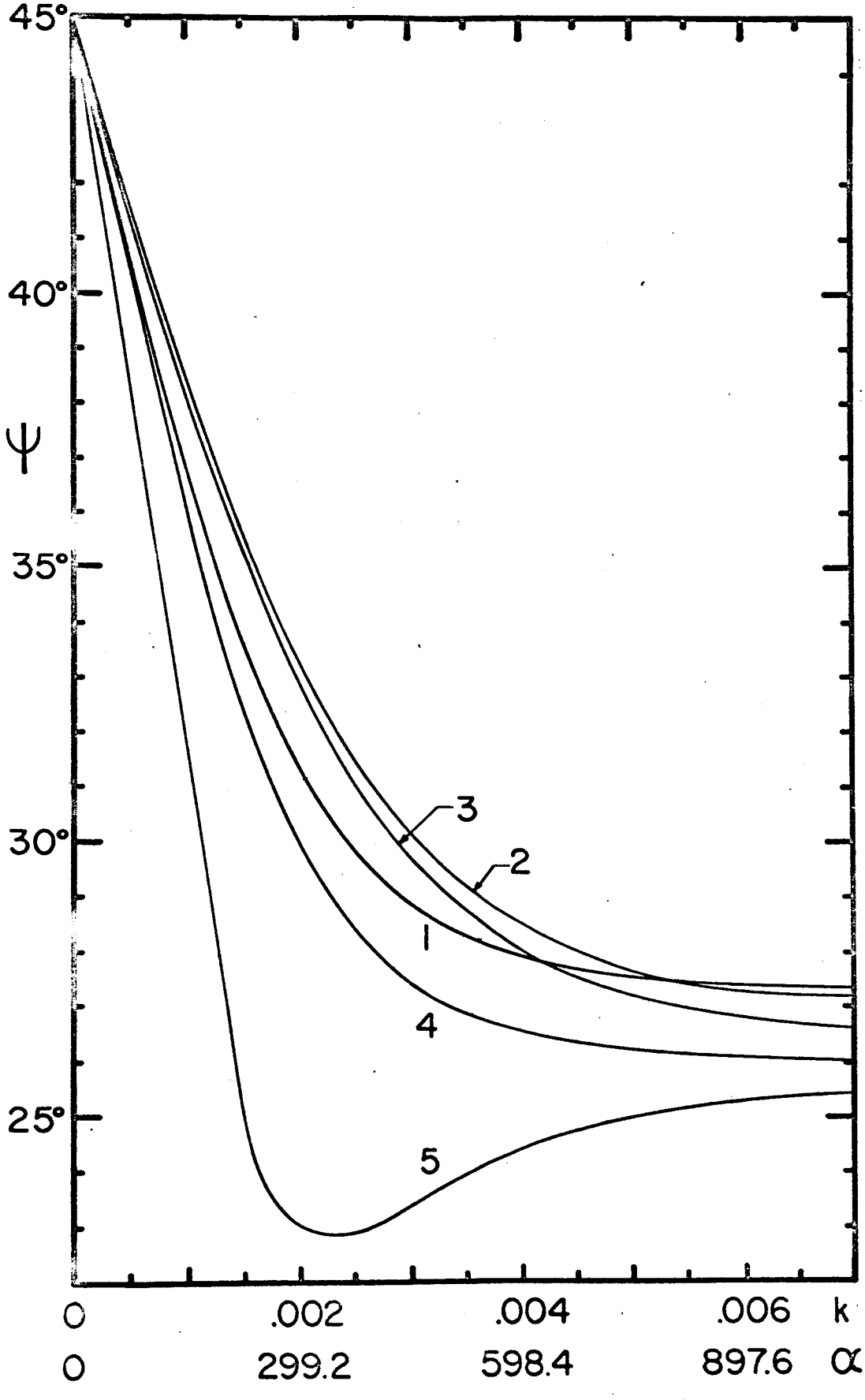


Fig. E-4 Plot of theoretical values of ψ (degrees) versus k (upper abscissa) and versus α (cm^{-1}, lower abscissa) for the three media geometry. The curves are numbered in the manner of Fig. E-3. The normal incidence absorption coefficient α is given by $\alpha = 4\pi k/\lambda$.

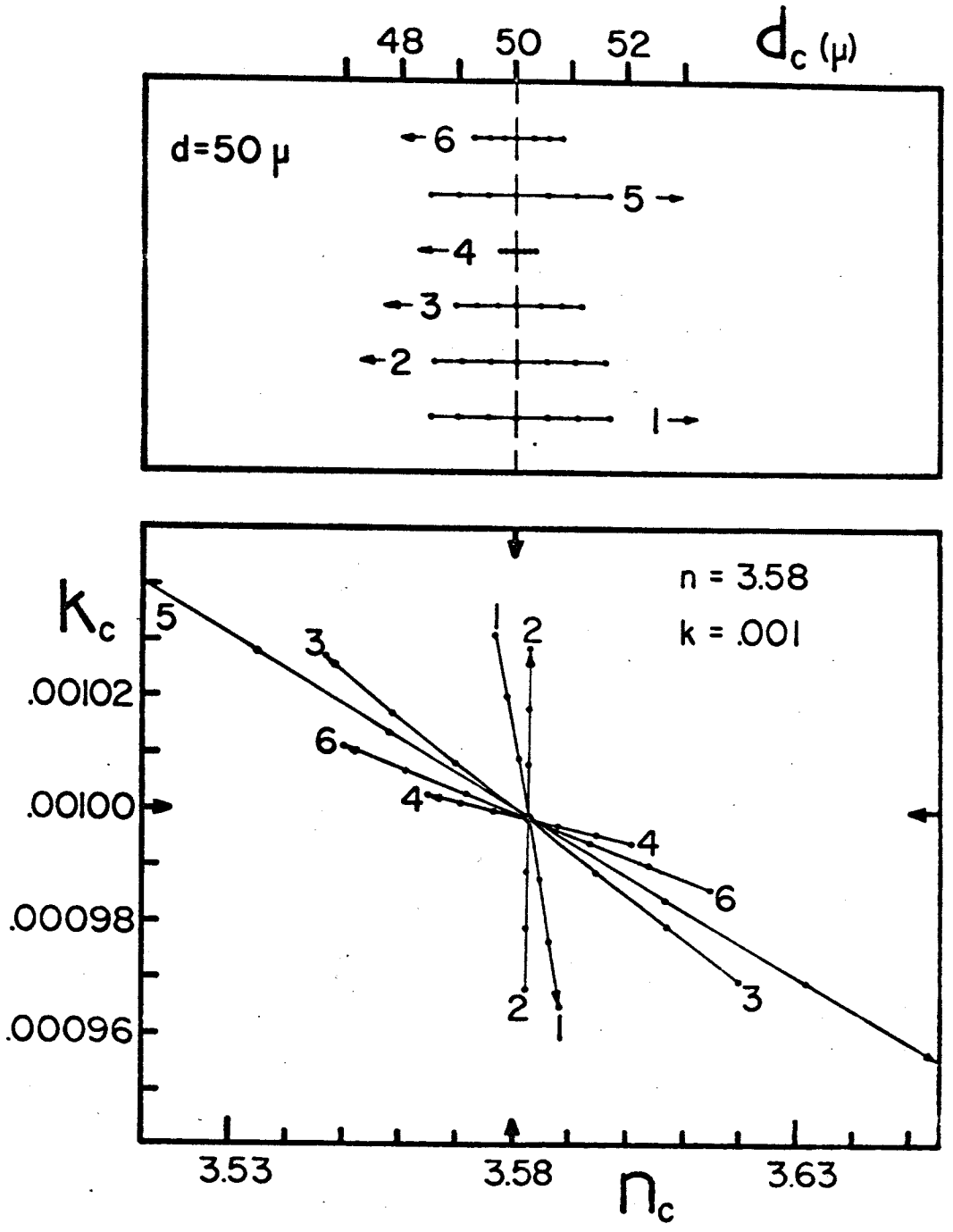


Fig. E-5 Graphs of calculated quantities d_c and $n_c + jk_c$ for the three media geometry with individual errors as the changing parameters. Each curve is drawn through seven calculated points indicated by dots (•), with the arrow indicating the direction of positive increase for the particular error. The numbering of the curves specifies which error is changing as well as its maximum value as follows. (1) $\delta i_1 = 0.06°$; (2) $\delta i_2 = 0.06°$; (3) $\delta \Delta_1 = 0.45°$; (4) $\delta \Delta_2 = 0.45°$; (5) $\delta \Psi_1 = 0.15°$; (6) $\delta \Psi_2 = 0.15°$. Thus curve 1 in the (n_c , k_c) plane is drawn through points $n_c + jk_c$ which correspond to $\delta i_1 = -0.06°, -0.04°, -0.02°, 0°, +0.02°, +0.04°, +0.06°$. The true values are $d = 50\mu$, $n = 3.58$, and $k = 0.001$.

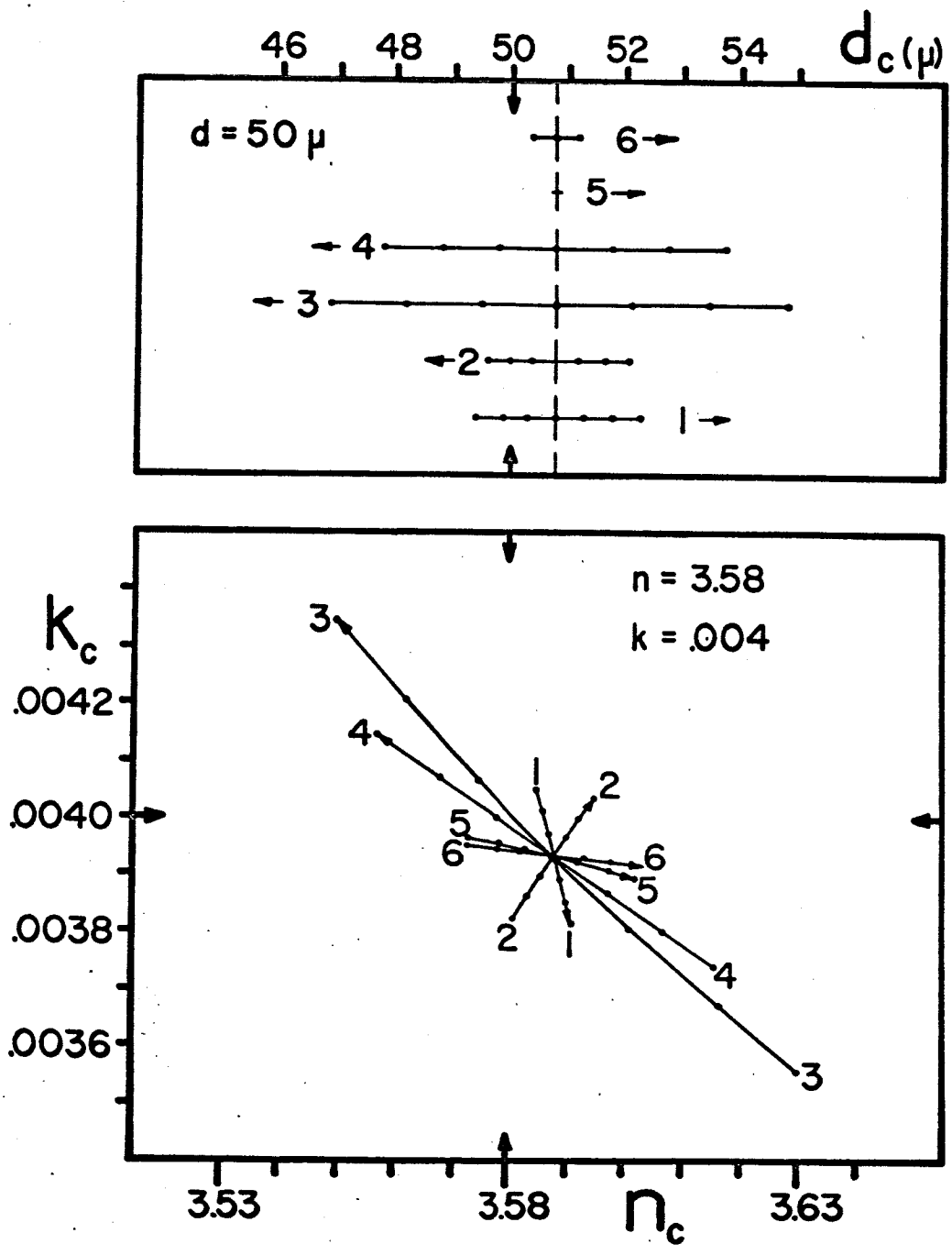


Fig. E-6 This figure presents the same type of data as does Fig. E-5 except that k = 0.004.

This would be the case, for example, if we happened to be interested in how k is affected by variations in types and concentrations of doping materials for the GaAs case we have examined.[1]

E.3. DISCUSSION AND CONCLUSIONS

We have examined two and three media techniques for studying the optical properties of an isotropic, homogeneous, weakly absorbing substance. If one could use perfectly collimated light and obtain perfect measurements, the two media approach would be preferred because $n_c + j k_c$ is then precisely equal to $n + j k$ while, for the three media case, $n_c + j k_c$, even for perfect measurements does not converge precisely to $n + j k$ because of the approximations used to derive the inversion formulae (32a,b,c). However, when one investigates the effects of realistic measurement errors, the three media approach is more attractive because no single measurement error is drastically amplified in the inversion equations.

For the three media case, as in the two media case, we have not presented an optimum method for selecting i_1 and i_2. We have shown, however, that an ellipsometric technique using a three media geometry can be used to gain estimates of the optical properties of a weakly absorbing substance with the individual measurement errors contributing typically less than 10% error in n_c, k_c, and d_c.

APPENDIX F

The Use of Retardation Plates in Ellipsometry

In this appendix I obtain a new formula, for the use of $\lambda/4$ or near $\lambda/4$ plates, with particular emphasis on its application in ellipsometric measurements. Consider the geometry of Fig. F-1. An electromagnetic wave is assumed to be traveling in the $+z$ direction, out of the paper. The x and y axes are reference axes. A retardation plate intercepts the wave such that the fast axis of the plate makes

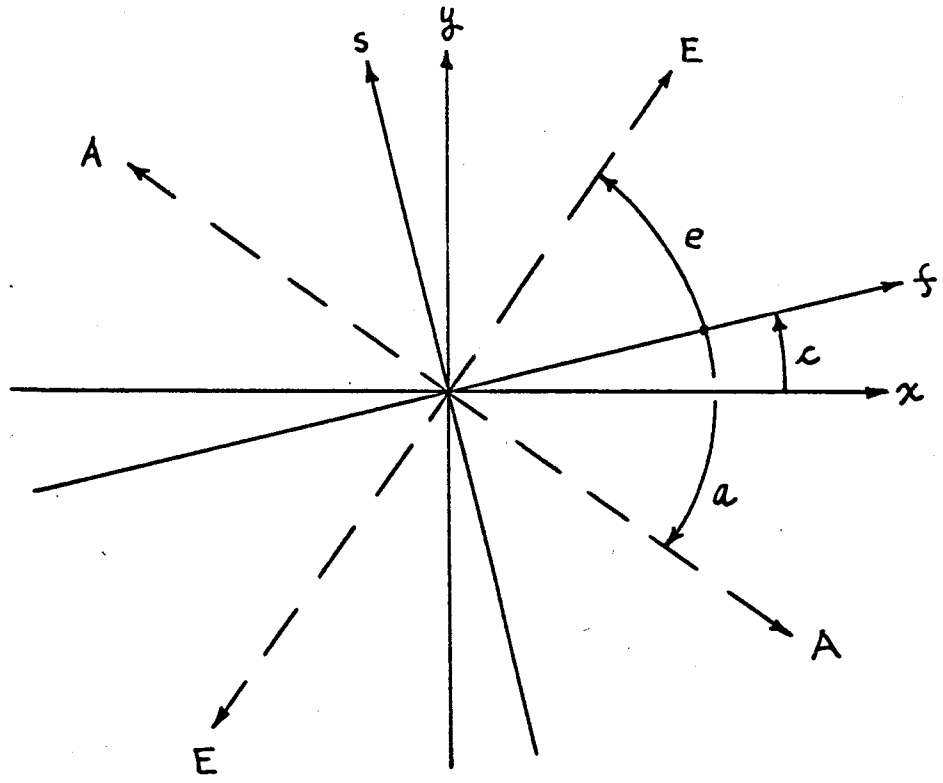


Fig. F-1. Geometry under consideration.

an angle c with the positive x axis. All physically distinct positions of the fast axis can be achieved by limiting c to the range $0 \leqq c < \pi$. In this appendix we shall define a positive fast axis and a positive slow axis as indicated by the f and s directions of Fig. F-1.

In using the retardation plate to measure the polarization state of the incident wave, the fast axis is rotated about the z axis until the wave emerging from the plate is linearly polarized. Suppose that the resultant linearly polarized electric field vector makes an angle e with respect to the positive fast axis. All physically distinct linear vibrations can be obtained if $-\pi/2 < e \leq \pi/2$. The emerging electric field can be extinguished by an analyzer if the transmission axis of the analyzer is perpendicular to the emerging electric field vector. The extinction position of the transmission axis is denoted by the angle a in Fig. F-1. Again, $-\pi/2 < a \leq \pi/2$. Note that the analyzer transmission axis is located with respect to the positive fast axis of the compensator.

The incident electric field can be decomposed into the components

$$\mathcal{E}_{ix} = E_{ix} \exp\{j\omega t - j2\pi z/\lambda\},$$
$$\mathcal{E}_{iy} = E_{iy} \exp\{j\omega t - j2\pi z/\lambda\}, \tag{1F}$$

where it is assumed that an air environment surrounds the retardation plate and the analyzer. The polarization state of the incident wave is defined by the complex number ρ, where

$$E_{iy}/E_{ix} = \tan\psi \exp j\Delta \equiv \rho. \tag{2F}$$

In (2F), the angles are restricted to the ranges $0 \leq \psi \leq \pi/2$ and $0 \leq \Delta < 2\pi$. It is important to note that the polarization state of the incident wavehas been defined with respect to the xy reference axes. Equations of the form (2F) frequently arise in ellipsometry and it would be very convenient if the complex number ρ could be measured directly. In using a compensator-analyzer combination, the measured quantities are the angles c and a, hence, the next best thing would be to find an explicit relation for ρ as a function of c and a. The explicit solution for ρ will be given in this appendix.

If the positive fast axis of the compensator is at an angle c, then a coordinate transformation yields

$$E_{is}/E_{if} = (\rho - \tan c)/(1 + \rho \tan c) , \tag{3F}$$

where E_{is} (E_{if}) is the complex amplitude of the incident electric field component along the positive slow (fast) axis. Analogously to (2F), the polarization state of the incident wave, with respect to the fs reference frame, is specified by

$$E_{is}/E_{if} = \tan\varphi \exp j\delta , \tag{4F}$$

where $0 \leqq \varphi \leqq \pi/2$ and $0 \leqq \delta < 2\pi$.

The output wave emerging from the compensator has f and s electric field components characterized by

$$E_{os}/E_{of} = T \exp(-j\Delta_c)[E_{is}/E_{if}] . \tag{5F}$$

As shown in App. B, the geometrical optics approximation leads to

$$T = 1 \quad ; \quad -\Delta_c = (2\pi d/\lambda)(n_f - n_s) , \tag{6F}$$

where d is the plate thickness. Since $n_s > n_f$, by definition, the quantity Δ_c is a positive number. Note that Δ_c and T are fixed quantities and are therefore independent of the rotational setting of the fast axis of the compensator plate. If the plate is absorbing and can be characterized by complex refractive indices of the form $n_s - jk_s$ and $n_f - jk_f$, then the exact or electromagnetic solution for Δ_c and T is given by

$$T \exp(-j\Delta_c) = \{\cosh\Gamma_f + K_f \sinh\Gamma_f\}/\{\cosh\Gamma_s + K_s \sinh\Gamma_s\} ,$$

$$\Gamma_m = j2\pi d(n_m - jk_m)/\lambda , \quad K_m = \frac{(n_m - jk_m)^2 + 1}{2(n_m - jk_m)} , \quad m = f, s .$$

Since the purpose of the compensator is to produce a linearly polarized output wave, then at some position C we must have

$$\delta - \Delta_c = m\pi \ , \quad m = 0, \pm 1, \pm 2, \ldots \tag{7F}$$

By substituting (4F) and (2F) into (3F) we find that

$$\delta = \tan^{-1}\{\sin\Delta/(\cos 2c \cos\Delta - \sin 2c \cot 2\psi)\} \ , \tag{8F}$$

$$\tan\varphi = \left[\frac{\tan^2 c + \tan^2\psi - 2\tan c \tan\psi \cos\Delta}{1 + \tan^2 c \tan^2\psi + 2\tan c \tan\psi \cos\Delta}\right] \ . \tag{9F}$$

Eqtn. (8F) has been obtained by Hall[35] in his paper on inexact quarter-wave plates. By writing (7F) as $\tan\delta = \tan\Delta_c$, then the solutions for c which specify the fast axis positions for compensation are found from (8F) to be given by $\tan c =$

$$\frac{-\cot 2\psi \sin\Delta_c \pm \left[\cot^2 2\psi \sin^2\Delta_c - \sin(\Delta+\Delta_c)\sin(\Delta-\Delta_c)\right]^{1/2}}{\sin(\Delta+\Delta_c)} \ . \tag{10F}$$

When c is between 0 and π , (10F) indicates that a linearly polarized output wave can be obtained for two distinct positions of the positive fast axis. If the quantity within the square root brackets in (10F) is not positive, then a linearly polarized output wave cannot be achieved. When the compensator is an exact quarter-wave plate, i.e., $\Delta_c = \pi/2$, then the two solutions for c differ by 90°.

The vibration direction of the resultant linearly polarized electric field is specified by the angle e where

$$\tan e = E_{os} / E_{of} \ . \tag{11F}$$

If the linear vibration is extinguished by an analyzer, then the position of the transmission axis is specified by the angle a where

$$-\cot a = \tan e \ . \tag{12F}$$

Note that, when e is positive, a is negative and vice-versa.

Suppose now that the compensation-extinction operations are carried out at the two possible compensator settings. The four measured angles are denoted by C_1 , a_1 , C_2 , and a_2. The relation between the measurements and the unknown ρ is

$$\frac{-\cot a_1}{-\cot a_2} = \frac{T\exp(-j\Delta_c)\cdot(\rho-\tan C_1)/(1+\rho\tan C_1)}{T\exp(-j\Delta_c)\cdot(\rho-\tan C_2)/(1+\rho\tan C_2)} \cdot \qquad (13F)$$

The factor $T\exp(-j\Delta_c)$ is common to both numerator and denominator in (13F) and can be cancelled. Note that the numerator (denominator) of the left side of (13F) is equal to the numerator (denominator) of the right side. The solution of (13F) for ρ is

$$\rho = \left\{ \cos(C_2+C_1)\sin(a_2-a_1) \pm j\left[\sin^2(C_2-C_1)\sin^2(a_2+a_1) - \sin^2(a_2-a_1)\right]^{1/2} \right\} \times \left\{ \sin(C_2-C_1)\sin(a_2+a_1) - \sin(C_2+C_1)\sin(a_2-a_1) \right\}^{-1} . \qquad (14F)$$

I have not yet determined whether the + or - sign in (14F) is appropriate. The correct sign culd be determined by further analysis; however, since most ellipsometry data eventually winds up on a computer program, the correct sign in (14F) would probably most easily be determined by working out some numerical examples on the computer.

Eqtn. (14F) represents the significant advance of this appendix. Many advantages accure from having an explicit solution for ρ as a function of measured quantities. The phase difference Δ can be assigned to the correct quadrant, thus enabling one to determine without doubt whether the elliptic polarization is left-handed or right-handed. In Chapt. 3 we found that expressions of the form (2F) could be used to specify the polarization state of the wave reflected from a specimen. By orienting the x axis in Fig. F-1 perpendicular to the

plane of incidence in an ellipsometry geometry, thequantity ρ_r characterizing the reflected wave could be directly computed by (14F). The usual procedure in ellipsometry work is to calculate first the direction of the major axis and the ellipticity of the elliptic vibration and then calculate Δ and Ψ . The equations used in the conventional procedure offer no rules for specifying whether the elliptic vibration is left or right-handed,[35] and, further, the numerical values for T and Δ_c must be known (See App. B).

Another advantage of (14F) is that nothing need be known about the specific characteristics of the retardation plate other than the direction of the fast axis. Whether the plate is exactly quarter-wave or whether multiple internal reflections are present become immaterial considerations when (14F) is used. If one is interested in calibrating the compensator, this can be done by inserting a calculated value of ρ into

$$-\cot a = T \exp(-j\Delta_c)\cdot(\rho - \tan c)/(1+\rho \tan c), \qquad (15F)$$

to obtain the calibration factor $T\exp(-j\Delta_c)$. Thus a calibration apparatus would consist of the compensator to be calibrated, an analyzer (polarizing device such as a Glan-Thompson prism), a detector, and a source of elliptically polarized, well collimated, monochromatic light of the desired wavelength. Several alternative calibration techniques, some of which are rather elaborate, are discussed by Jerrard[36].

The measurement procedure can be summarized as follows. The reference axes corresponding to the positive x and y directions in Fig. F-1 are chosen. The fast axis of the compensator is aligned parallel to the x axis. In this position, the positive y direction is specified as the positive x direction. The compensator and the analyzer are independently rotated until extinction occurs. When looking into the incident beam, the compensator angle c is positive when the positive fast axis is rotated counter-clockwise from the positive axis. Also, the analyzer angle a is positive when the transmission axis position is obtained by a counterclockwise rotation of the trans-

mission axis from the positive fast axis. In Fig. F-J, C is positive and a is negative. The angular measurements are recorded for the two possible compensation positions of the fast axis and the measurements are then substituted into (14F) to obtain the value of ρ appropriate to the xy reference frame. For the sake of consistency and to avoid any possible ambiguities, it is probably wise to use the restrictions $0 \leqq C < \pi$ and $-\pi/2 < a \leqq +\pi/2$.

APPENDIX G

Polarization State of Thin Film Reflection

This appendix is a relatively self contained work and has been published in the Journal of the Optical Society of America, vol. 55, p. 577 (1965). The differences between this appendix and the published article are only minor. Permission to reproduce the published work has been granted by the Journal of the Optical Society of America. The figures and figure captions relevant to this appendix can be found in numerical sequence at the end of the appendix.

Polarization State of Thin Film Reflection*

D. A. HOLMES AND D. L. FEUCHT

Department of Electrical Engineering
Carnegie Institute of Technology
Pittsburgh, Pennsylvania 15213

Hacskaylo[1] has described a rapid and accurate technique for determining the refractive index of a transparent film deposited on a transparent substrate. In his paper, Hacskaylo references Vasicek[2] in stating that, "The elliptically polarized light can be approximated as linearly polarized light by noting that the reflected light from a plane parallel homogeneous film is slightly elliptically polarized and the reflected light vector describes a narrow ellipse." While this statement is clearly not quantitative, it is, nontheless, quite general since Hacskaylo has not stated any conditions or restrictions. Because the polarization state of a monochromatic plane wave after reflection from a plane parallel film depends on many factors, including the polarization state of the incident wave, the refractive indices of film and substrate, the film thickness, the angle of incidence, and the wavelength, it is instructive to consider further the validity of approximating the elliptically polarized reflected light as linearly polarized light.

The electric field vector of the reflected wave can be decomposed into p and s components[3]. These components can be characterized by a quantity δ_r, defined as the phase difference between the p and s

---

*This work was supported by the National Aeronautics and Space Administration under Carnegie Institute of Technology Contract NAS8-5269.

[1] Michael Hacskaylo, J. Opt. Soc. Am. 54, 198 (1964).

[2] A. Vasicek, Optics of Thin Films, (North-Holland Publishing Co., Amsterdam, The Netherlands, 1960), p. 86.

[3] Except when specified otherwise, our geometry and notation are the same as those of Hacskaylo.

components, and a quantity $\tan\gamma_r$, defined as the ratio of the magnitude of the p component to the magnitude of the s component. The ellipticity e , defined as the ratio of the minor axis to the major axis of the reflected elliptic vibration, is given by[4]

$$e = \tan\left\{\frac{1}{2}\sin^{-1}(\sin\delta_r \sin 2\gamma_r)\right\} . \qquad (1)$$

The electric field vector of the incident wave can also be characterized by a phase difference δ_i and an amplitude ratio $\tan\gamma_i$, and, when this is done, we find that[5]

$$\tan\gamma_r = \tan\gamma \tan\gamma_i ,$$

$$\delta_r = \Delta + \delta_i , \qquad (2a)$$

$$\tan\gamma \cdot e^{j\Delta} \qquad (2b)$$

where the quantity is nothing more than the polar form of R_p/R_s. Now by using (2a) and (2b) in (1) we have

$$e = \tan\left\{\frac{1}{2}\sin^{-1}\left(\frac{2\sin[\Delta+\delta_i]\tan\gamma\tan\gamma_i}{1+\tan^2\gamma\tan^2\gamma_i}\right)\right\} . \qquad (3)$$

In order to simplify the interpretation of (3), let us assume that the incident wave is linearly polarized with $\delta_i = 0$. In this case, the angle γ_i can be thought of as the azimuth (with respect to the s axis) of the transmission axis of an input polarizer. The ellipticity e can be made small, perhaps negligible, by adjusting γ_i so that the product $\tan\gamma\tan\gamma_i$ is either very large or very small, the limiting case of $e=0$ being reached when either $\gamma_i = 0^\circ$ or $\gamma_i = 90^\circ$. When Δ is near $m\pi$, $m = 0, \pm1, \pm2, \ldots$, we see, from (3), that e

---

[4] M. Born and E. Wolf, Principles of Optics, (Pergamon Press, London, England, 1959), pp. 26-27. The x and y axes defined in this reference coincide, respectively, with the s and p axes for the reflected wave. In (1), the argument of the tangent function must be chosen between -45° and $+45^\circ$ in order that $-1 \leq e \leq +1$.

[5] A more detailed development of equations similar to (2a) and (2b) is contained in D.A. Holmes, J. Opt. Soc. Am. 54, 1340(1964). Equations (2a) and (2b) relate the polarization state of the reflected wave to the polarization state of the incident wave and, through the quantities Δ and γ, show how the polarization state of the reflected wave is influenced by the thin film reflection. Note that Δ and γ are independent of δ_i and γ_i .

will be near zero, regardless of the value of $\tan\gamma_i$, however, when Δ is near $(2m+1)\pi/2$, $|e|$ can approach unity by adjusting γ_i so that $\tan\gamma\tan\gamma_i \cong 1$. A circularly polarized reflected wave can be achieved if Δ is an odd integral multiple of π and if the incident wave is polarized such that $\tan\gamma\tan\gamma_i = 1$.

Some other features of dielectric film reflection can best be illustrated by a numerical example. Let us consider a case discussed by Hacskaylo, namely, reflection from a CaF_2 film ($n_1 = 1.2174$) deposited on a microscope slide ($n_2 = 1.5108$). For some selected film thicknesses, Fig. G-1a shows thetheoretical variation of Δ with the angle of incidence ϕ_o , while Fig. G-1b shows thetheoretical variation of $\tan\gamma$ with ϕ_o . In the numerical calculations[6], we set $R_p/R_s = u + jv$ and then, by calculating u and v , were able to assign Δ to the correct quadrant. Note that, for the sake of showing Δ as a continuous function of ϕ_o , we have <u>not</u> restricted Δ to the range $0 \leqq \Delta \leqq 2\pi$. For all of the film thicknesses, circularly polarized light can be achieved; however, for the very thin film (curve 1), $\tan\gamma_i$ is required to be about 20 or $\gamma_i \cong 87^\circ$. An important feature shown by the figures is that the value of ϕ_o , for which circularly polarized reflected light is possible, can vary over tens of degrees, depending on the value of film thickness.

From the above discussion, we conclude that an elliptically polarized wave, obtained by reflecting a linearly polarized wave from a lossless film on a lossless substrate, cannot always be approximated as linearly polarized.

We wish to thank Professor R. L. Longini, Carnegie Institute of Technology, for several helpful suggestions.

---

[6] We have used the sign convention for the Fresnel formulae used in ref. 4, p. 39.

FIGURE CAPTION FOR APPENDIX G

Fig. G-1 Parameters which describe the polarization state of reflections from a CaF_2 film deposited on a microscope slide, assuming refractive index values reported by Hacskaylo. (a) Theoretical values of Δ (degrees) vs. angle of incidence ϕ_0 (degrees). (b) Theoretical values of $\tan\gamma$ vs. ϕ_0 (degrees). The curve numbers specify the film thickness d as follows: (1) $d = \lambda/20$, (2) $d = \lambda/5$, (3) $d = \lambda/4$, and (4) $d = 3\lambda/10$, where λ is the vacuum wavelength. The curves were plotted from computer calculations of Δ and $\tan\gamma$, using $\phi_0 = 0^\circ, 2^\circ, 4^\circ, \ldots, 90^\circ$.

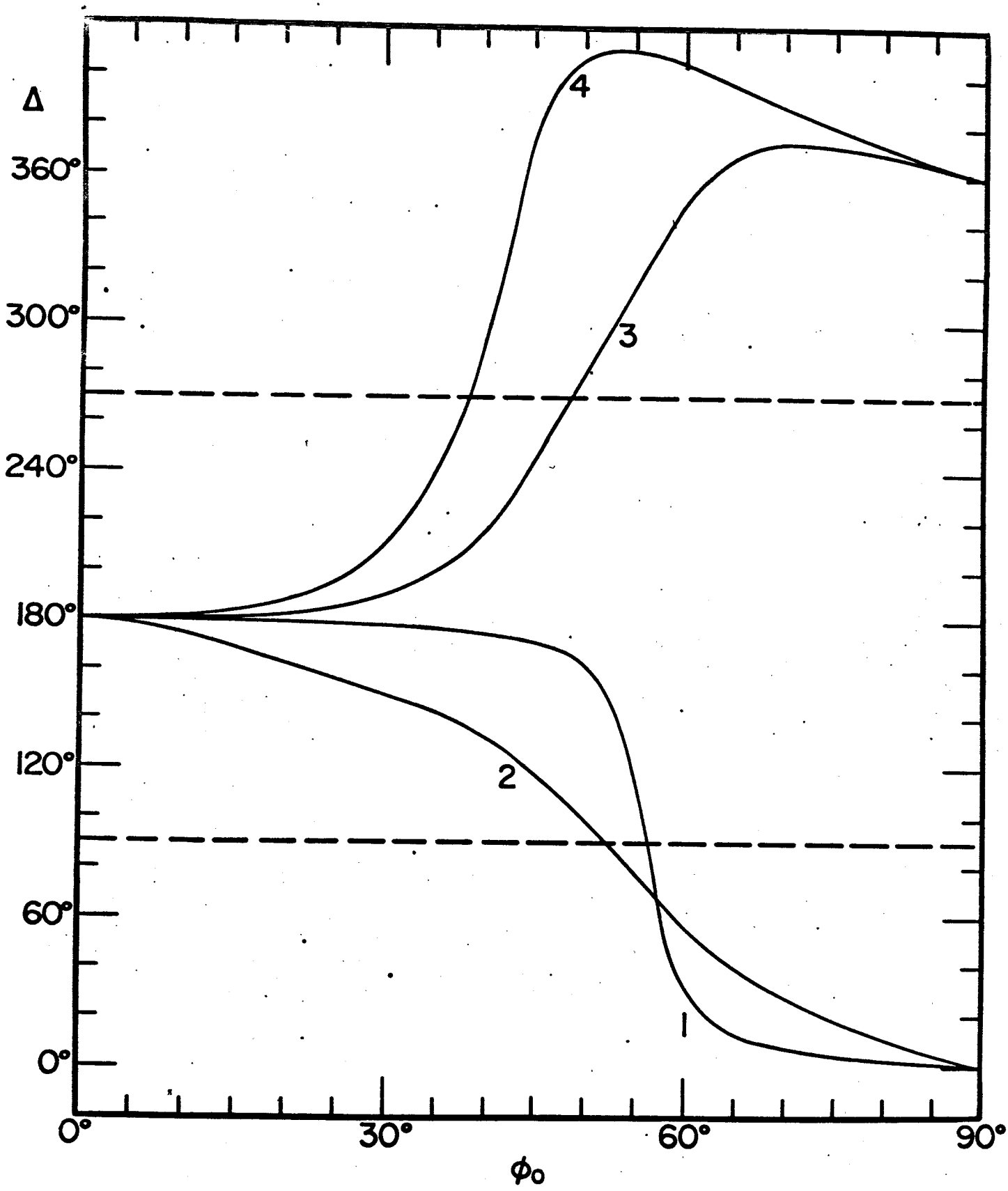


Fig. G-1a

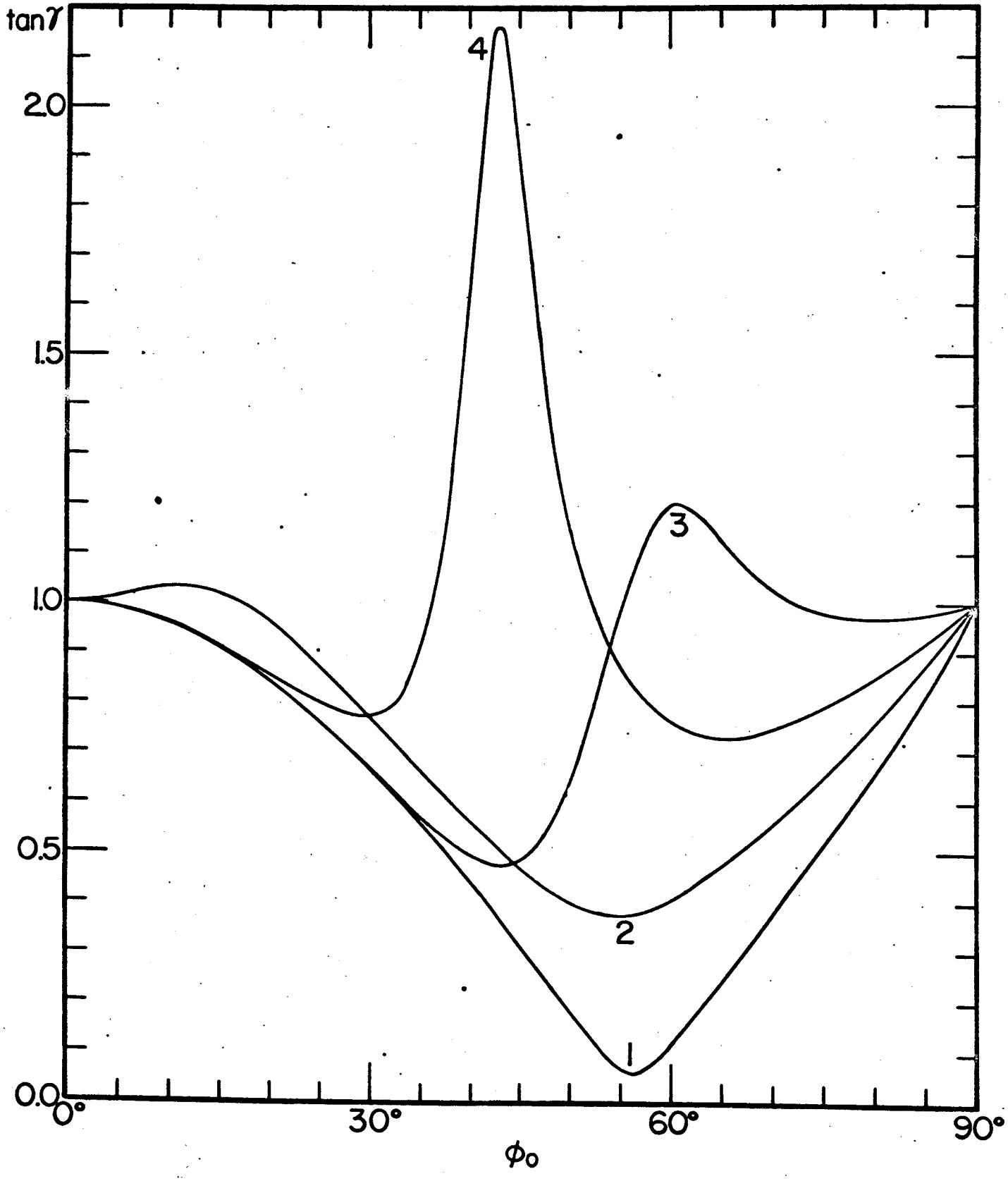


Fig. G-1b

REFERENCES

1. G. N. Ramachandran and S. Ramaseshan in Handbuch der Physik, edited by S. Flugge, (Springer-Verlag, Berlin, 1961), Vol. 25., Chapt. 1, p. 111.

2. H. Schopper, Z. Physik 132, 146 (1952).

3. A. B. Winterbottom, Kgl. Norske Videnskab. Selskabs Skrifter 1, 27, 37 (1955).

4. A. M. Goncharenko and F, L. Federov, Opt. Spectry. 13, 48 (1962).

5. M. Born and E. Wolf, Principles of Optics (Pergamon Press, Inc., New York, 1959), pp. 54, 61.

6. Ref. 5, p. 705.

7. B. I. Stepanov and A. P. Khapalyuk, Opt. Spectry. 13, 404 (1962).

8. W. A. Shurcliff and S. S. Ballard, Polarized Light (D. Van Nostrand Co., Inc., Princeton, 1964), pp. 42-54.

9. H. A. Daw, J. Opt. Soc. Am. 53, 915 (1963).

10. E. L. Steele, W. C. Davis, and R. Treuthart, J. Opt. Soc. Am. 54, 1399A (1964).

11. R. G. Fellers and J. Taylor, IEEE Trans. MIT MIT-12, 584 (1964).

12. H. Jacobs, D. A. Holmes, L. Hatkin, and F. A. Brand, J. Opt. Soc. Am. 54, 1416 (1964); IEEE Trans. MIT MIT-12, 163 (1964); J. Appl. Phys. 34, 2617 (1963).

13. R. E. Collin, Field Theory of Guided Waves (McGraw-Hill Book Co., Inc., New York, 1960), pp. 89-90.

14. J. A. Mandarino, Am. Mineralogist 44, 65 (1959).

15. A. K. Levine, Am. Scientist 51, 14 (March, 1963).

16. Ref. 5, pp. 688-691.

17. W. A. Shurcliff, Polarized Light (Harvard University Press, Cambridge, 1962), pp. 87-108.

18. H. G. Jerrard, J. Opt. Soc. Am. 38, 35 (1948).

19. W. J. Turner and W. E. Reese, J. Appl. Phys. 35, 350 (1964).

20. D. T. F. Marple, J. Appl. Phys. 35, 1241 (1964).

21. R. J. Archer, Phys. Rev. 110, 354 (1958).

22. R. C. Menard, J. Opt. Soc. Am. 52, 427 (1962).

23. Charles D. Hodgman, Editor in Chief, Handbook of Chemistry and Physics (Chemical Rubber Publishing Co., Cleveland, 37th edition, 1955-56), p. 2682.

24. K. S. Champlin and D. B. Armstrong, IEEE Trans. MIT. MIT-11, 73 (1963). R. T. Jacobson, J. Opt. Soc. Am. 54, 1170 (1964).

25. D. K. Burge and H. E. Bennett, J. Opt. Soc. Am. 54, 1428 (1964).

26. E. G. Coker and L. N. G. Filon, A. Treatise on Photoelasticity (Cambridge University Press, London, 1957).

27. E. Poindexter, Am. Mineralogist 40, 1032 (1955).

28. A. A. Giardini, J. Opt. Soc. Am. 47, 726 (1957).

29. Ref. 1, pp. 202-217.

30. Ref. 5, pp. 700-702.

31. C. Hilsum and A. C. Rose-Inness, Semiconducting III-V Compounds (Pergamon Press, London, 1961), p. 7.

32. Ref. 1, pp. 85-96.

33. F. Partovi, J. Opt. Soc. Am. 52, 918 (1962).

34. Ref. 13, pp. 97-100.

35. A. C. Hall, J. Opt. Soc. Am. 53, 801 (1963).

36. H. G. Jerrard, J. Opt. Soc. Am. 42, 159 (1952).

37. M. D. Sturge, Phys. Rev. 127, 768 (1962).

38. C. A. Hampel, editor, Rare Metals Handbook (Reinhold Publishing Corp., London, 2nd ed., 1961), pp. 672-673.

39. F. B. Seely and J. O. Smith, Resistance of Materials (John Wiley and Sons, Inc., New York, 4th ed., 1956) pp. 267-270 and 380-381.

40. G. Burns and M. I. Nathan, Proc. IEEE 52, 770 (1964).

Bibliography

ELLIPSOMETRIC MEASURING TECHNIQUES

R. J. Archer, "Optical Measurement of Film Growth on Silicon and Germanium Surfaces" J. Electrochem. Soc. 104, 619 (1957).

R. J. Archer, "Optical Constants of Germanium, "Phys. Rev. 110, 354 (1958).

R. J. Archer,"Determination of Properties of Films on Silicon by Method of Ellipsometry," J. Opt. Soc. Am. 52, 970 (1962).

J. B. Bateman and M. W. Harris, "Measurement of the Properties of Thin Films on Chromium by the Reflection of Polarized Light", Ann. New York Aca. Sci. 53, 1064 (1951).

D. Bergman, "Analysis of Elliptically Polarized Light", J. Opt. Soc. Am. 52, 1080 (1962).

D. K. Burge and H. E. Bennett, "Effect of a Thin Surface Film on the Ellipsometric Determination of Optical Constants", J. Opt. Soc. Am. 54, 1428 (1964).

E. N. Cameron, "Apparatus and Techniques for the Measurement of Certain Optical Properties of Ore Minerals in Reflected Light", Econ. Geol. 52, 252 (1957).

R. W. Ditchburn and G. A. J. Orchard, "The Polarization of Totally Reflected Light," Proc. Physical Soc. LXVII, 608 (1954).

R. W. Ditchburn, "Some New Formulas for Determining the Optical Constants from Measurements on Reflected Light", J. Opt. Soc. Am. 45, 743 (1955).

J. A. Faucher, G. M. McManus, and H. J. Trurnit, "Simplified Treatment of Ellipsometry", J. Opt. Soc. Am. 48, 51 (1958).

F. L. McCrackin, E. Passaglia, R. R. Stromberg, and H. L. Steinberg, "Measurement of the Thickness and Refractive Index of Very Thin Films and the Optical Properties of Surfaces by Ellipsometry," J. Res. Natl. Bur. Stds. 67A, 363 (1963).

F. L. McCrackin and J. P. Colson, "A Fortran Program for Analysis of Ellipsometer Measurements and Calculation of Reflection Coefficients from Thin Films", Natl. Bur. Stds. Tech. Note 242, (27 May 1964).

R. C. Menard, "Optical Measurement of Oxide Thickness on Titanium," J. Opt. Soc. Am. 52, 427 (1962).

F. P. Mertens, P. Theroux, and R. C. Plumb, "Some Observations on the Use of Elliptically Polarized Light to Study Metal Surfaces," J. Opt. Soc. Am. 53, 788 (1963).

F. Partovi, "Theoretical Treatment of Ellipsometry", J. Opt. Soc. Am. 52, 918 (1962).

R. L. Patrick, D. A. Ross, and W. A. Vaughan, "Ellipsometry", U. S. Dept. of Commerce Doc. PB 121663, (March, 1956).

R. F. Potter, "Analytical Determination of Optical Constants Based on the Polarized Reflectance at a Dielectric-Conductor Interface", J. Opt. Soc. Am. 54, 904 (1964).

A. Rothen and M. Hanson, "Optical Measurements of Surface Films. I," Rev. Sci. Instr. 19, 839 (1948).

A. Rothen, "The Ellipsometer, an Apparatus to Measure Thickness of Thin Surface Films", Rev. Sci. Instr. 16, 26 (1945).

A. Rothen and M. Hanson, "Optical Properties of Surface Films. II," Rev. Sci. Instr. 20, 66 (1949).

A. Rothen, "Optical Properties of Surface Films", Ann. New York Aca. Sci. 53, 1054 (1951).

A Rothen, "Improved Method to Measure the Thickness of Thin Films with a Photoelectric Ellipsometer", Rev. Sci. Instr. 28, 283 (1957).

R. R. Stromberg, E. Passaglia, and D. J. Tutas, "Thickness of Adsorbed Polystyrene Layers by Ellipsometry", J. Res. Natl. Bur. Stds. 67A, 431 (1963).

A. B. Winterbottom, "Optical Methods of Studying Films on Reflecting Bases Depending on Polarization and Interference Phenomena," Trans. Faraday Soc. 42, 487 (1946).

A. B. Winterbottom, "Optical Studies of Metal Surfaces; Especially on the Use of the Reflected-Polarized-Light Method in Investigating Surface Films", Det Kgl Norske Videnskab. Selskabs. Skrifter 1, 1 (1955).

K. H. Zaininger and A. G. Revesz, "Ellipsometry-A Valuable Tool in Surface Research", RCA Rev. 25, 85 (1964).

STRAIN AND PRESSURE EFFECTS IN GaAs LASERS

J. Feinleib, S. Groves, W. Paul and R. Zallen, "Effect of Pressure on the Spontaneous and Stimulated Emission from GaAs," Phys. Rev. 131, 2070 (1963).

G. E. Fenner, "Effect of Hydrostatic Pressure on Emission from Gallium Arsenide Lasers", J. Appl. Phys. 34, 2955 (1963).

T. A. Fulton and D. B. Fitchen, "Pressure Effects in Ga ($As_{1-x}P_x$) Electroluminescent Diodes", Appl. Phys. Lett. 4, 9 (1964).

D. Meyerhofer and R. Braunstein, "Frequency Tuning of GaAs Laser Diode by Uniaxial Stress", Appl. Phys. Lett. 3, 171 (1963).

F. M. Ryan and R. C. Miller, "The Effect of Unizxial Strain on GaAs Laser Diodes", Appl. Phys. Lett. 3, 162 (1963).

M. J. Stevenson, J. D. Axe and J. R. Lankard, "Line Widths and Pressure Shifts in Mode Structure of Stimulated Emission from GaAs Junctions", IBM J. Res. Dev. 7, 155 (1963).

P-N JUNCTION LASERS

G. Burns and M. I. Nathan, "P-N Junction Lasers", Proc. IEEE, 52, 770 (July, 1964).

R. N. Hall, "Coherent Light Emission from P-N Junctions", Solid-State Electronics 6, 405 (1963).

R. N. Hall, "Semiconductor P-N Junction Lasers", IEEE Student J. 1, 13 (May, 1963).

J. M. Lavine, D. M. Warschauer, T. Deutsch, A. A. Iannini, R. C. Ellis, and E. Adler, "GaAs Injection Lasers", Raytheon Electronic Progress VII, 17 (Jan.-Feb., 1963).

B. Lax, "Semiconductor Lasers", Science 141, 1247 (Sept. 27, 1963).

B. Lax, "Semiconductor Diode Lasers", Solid-State Design 4, 26 (Nov. 1964).

H. T. Mindon, "Gallium Arsenide Electroluminescent and Laser Diodes", Semi. Pro. and S. S. Technology 6, 34 (Aug. 1964).

T. M. Quist, "Semiconductor Lasers", Intern'l Science and Technology, 80 (Feb. 1964).

R. H. Rediker, "Infrared and Visible Light Emission from Forward-Biased P-N Junctions", Solid/State/Design 4, 19 (Aug. 1963).

R. H. Rediker, "Semiconductor Lasers", Physics Today 18, 42 (Feb. 1965).