402

PHOTOPRODUCTION EXPERIMENT AT FERMILAB

B. Knapp, W. Lee, L. Moscoso, P. Sokolsky

Columbia University, New York, N.Y.

G. Gladding, M. Gormley, R. Messner, T. O'Halloran, A. Wattenberg

University of Illinois, Urbana, Illinois

M. Binkeley, I. Gaines, R. Orr, J. Peoples, L. Read

Fermi National Laboratory, Batavia, Illinois

Spokesman: Wonyong Lee
(914) 591-8100

May 1975

CONTENTS

	Page
Introduction	1
I. Physics Objectives	2
II. Linearly Polarized Wide Band Photon Beam	4
III. Detector Improvement	5
A. Recoil Proton Detector	6
B. Cerenkov Detectors	6
C. π^0/γ Detector	7
IV. Request for Running Time	9

Appendix I: Calculation on Degree of Polarization Expected to Attain

Appendix II: Friedberg, R.; PR A10, 21 (1974).

INTRODUCTION

We have recently completed a photoproduction experiment and an experiment on dimuon production by neutrons at 300 and 380 GeV incident proton energies. In the neutron experiment, we observed inclusive ρ and $\psi(3100)$ production in addition to a measurement of the dimuon continuum. In the photoproduction experiment, we detected the $\psi(3100)$ and $\psi'(3700)$ and measured properties of the $\psi(3100)$. We are currently searching for new resonances in both hadronic and semi-leptonic decay modes with and without strange particles in the final states. We are also studying diffractively photoproduced multihadron final states. In these reactions, we have observed the $\rho \to \pi^+\pi^-$, $\rho' \to \pi^+\pi^-\pi^+\pi^-$, and $\omega \to \pi^+\pi^-\pi^0$, and a possible hint of $\rho'' \to \pi^+\pi^-\pi^+\pi^-\pi^+\pi^-$ and $\omega' \to \pi^+\pi^-\pi^+\pi^-\pi^0$. We also observed the $\varphi \to K^+K^-$ but not $\varphi' \to \pi^+\pi^-\phi(K^+K^-)$.

Although our analysis has not been completed, we can clearly see ways of improving our experiment to explore further exciting physics. We propose to carry out a photoproduction experiment which is a natural extension of E-87A but using considerably improved detectors. We also propose to linearly polarize the photon beam and study the possibility of circular polarization.

In Sec. I, we discuss the physics objectives of this proposal. In Sec. II, we give a brief summary of the

characteristics of the linearly polarized photon beam. Additional details on the calculation of the degree of polarization we expect to attain are contained in Appendix I. Appendix II is an article by R. Friedberg which discusses the optimum manner of circularly polarizing the photon beam. Section III outlines the improvements we wish to make to our present detector and the support we are requesting from Fermilab. In Sec. IV, we present our request for the beam time required for the physics objectives of this proposal.

I. PHYSICS OBJECTIVES

The physics objectives of this proposed experiment are very similar to the ones we have been pursuing but using considerably improved detectors. The objectives are:

i) Explore the mass range from 1-6 GeV for the presence of 1^- states in the multihadron and 2 lepton final states;

ii) Measure the cross section for diffractively photoproduced multihadron states with 20-300 GeV photons;

iii) Determine the existence or non-existence of particles, such as heavy leptons, charm mesons and baryons.

The most significant addition to our present apparatus will be the addition of two gas Cerenkov counters which will be used in a conventional manner to provide particle identification of the hadrons in the photoproduced final state.

In 87A we were able to identify K^o, Λ^o and $\bar{\Lambda}^o$ via their decay near the production vertex. We have detected 50 K V^o events. With the addition of the Cerenkov detectors, we will be able to identify $K^{\pm}/\pi^{\pm}/p^{\pm}$ particles in the photoproduced final states. We will have a large sample of events (a few million events) allowing us to examine in detail the composition of the photoproduced final states.

Improvement of the π^o/γ detector over the existing apparatus will enable us to completely identify and reconstruct very complicated topologies. Although at present we measure the energies of photons well, their directions are measured very inefficiently. Improved reconstruction of π^o/γ and detailed identification of all charged particles greatly improve detection and reconstruction of the decays of the higher mass resonances, clearly one of the more important aspects of our proposal.

The addition of a cryogenic target and recoil detector will enable us to determine the slope of the diffractive peak for $\psi(3.1)$ photoproduction in the reaction $\gamma + p \rightarrow \psi + p$ and $\gamma + d \rightarrow \psi + d$. Because the value of the slope has been measured to be small, $b \approx 1\text{-}3/\text{GeV}^2$, it is crucial to separate elastic events from the inelastic ones. Furthermore, the study of the inelastic final states may by itself be very interesting.

The technique for producing a linearly polarized photon beam using an anisotropic crystal is well known and has been developed and tested recently at Cornell and SLAC. The degree of linear polarization increases with photon energy. It will be very appealing to use this high degree of linear polarization as an aid in understanding the properties of the high energy photoproduction of 1^- particles, in addition to searching for new particles directly coupled to the photon.

The calculation using available crystals shows that we will get a disappointing amount of circular polarization due to the mosaic spread of the crystal. We are currently looking for crystals with a better mosaic spread in our effort to produce the circularly polarized photons. As we gain experience with the linearly polarized beam, it may be possible to obtain circularly polarized beam at Fermilab. With a circularly polarized beam it will be possible to confront one of the most interesting questions in physics - whether or not there exist weakly coupled, parity-violating neutral currents which <u>interfere</u> with the electromagnetic current.

II. LINEARLY POLARIZED WIDE BAND PHOTON BEAM

We have studied the possibility of linearly polarizing the broad band photon beam. A technique used successfully at lower energies should be even more efficient at Fermilab energies. By inserting highly oriented graphite in the beam

we can preferentially absorb a particular linear polarization through coherent production of electron-positron pairs.

Linear polarization of ~ 25% over most of the useful photon energy spectrum can be achieved with about 6 inches of graphite and an overall attenuation of 5-7. If it were desired to achieve polarizations of ~ 50%, this can be achieved with an overall attenuation of ~ 20.

The polarizer would be placed in a magnetic field of ~ 10 kG to deflect the electron-positron pair that would radiate soft photons in the graphite. The angle between the graphite planes and the photon beam direction must be maintained to within a few tenths of a milliradian. This can be achieved by mounting the crystals as a single rigid unit pinned at one end and driven by a micrometer at the other end. Power dissipation in the polarizer will be negligible and no cooling system will be required.

III. DETECTOR IMPROVEMENT

The detector system which will allow us to carry out the physics program described above is basically an extension and improvement of the magnetic spectrometer and particle identification system used in expt. E-87. We describe here the modifications and additions we plan to make to the existing experimental layout and we summarize how these detectors will extend and improve the scope of the physics program we will pursue.

As an illustration, Fig. 1 shows the detector layout we would use to study the diffractively, photoproduced hadronic final states. The new detectors shown in Fig. 2 consist of: a liquid (LH2/LD2) target together with a recoil proton detector, a pair of two multi-cell, gas Cerenkov counters and an improved π^o/γ detector.

A. Recoil Proton Detector

The detector consists of two modules, located on either side of the cryogenic target. Each module subtends a solid angle of $\Delta\phi = \pm 60^o$ and consists of two drift chambers and a set of dE/dX and time of flight counters.

The specific manner in which the recoil detector would be employed depends upon which phase of the physics program described earlier we are pursuing. For example, to determine the slope of the diffractive peak for $\psi(3.1)$ photoproduction on hydrogen and deuterium, the momentum transfer determined by the recoil detector, will be compared to that determined by the momentum measurements of the forward going decay products of the $\psi(3.1)$. A comparison of the two independent measurements of the momentum transfer will not only allow us to separate elastic events but also to identify and study inelastic ones.

B. Cerenkov Detectors

The two Cerenkov counters shown in Fig. 2 will be used in a conventional fashion to provide detailed information and particle identification on the character

of the photoproduced final state. A detailed design and test program is currently in progress to optimize the parameters of these counters for the range of secondary momenta contained by our spectrometer. The precise momentum range over which these counters will provide complete separation of pions, kaons, and protons, will not be determined until the final design is complete. We estimate that we will be able to achieve complete π-K-p separation up to $\sim$ 100 GeV/c.

C. π^0/γ Detector

The addition of a π^0/γ detector to the experimental apparatus described above will provide us with the ability to do both complete identification and reconstruction of very complicated final state topologies. The existing electromagnetic shower detectors already yield a measure of the energies of both photons and electrons with a precision on $\pm$ 4% at 16 GeV. We plan to add a new-type of proportional wire chamber that will measure the position of the electromagnetic shower after a depth of 4-6 radiation lengths into the electromagnetic cascade. This depth is large enough to ensure that virtually 100% of the photons have begun to cascade but small enough so that the center of gravity of the shower has not begun to spread significantly in a lateral direction.

The novel feature of the proportional wire chamber we will use to measure the photon conversion point is that it will measure the pulse height and the center of gravity of the electromagnetic cascade with an accuracy of (0.5-1.0) inches.

We have constructed a prototype proportional wire chamber of the type we plan to use. The signals induced on the cathode plane of the chamber, by the charged particles in the electromagnetic cascade of a converted photon, are detected and pulse height analyzed. We have just completed the development of the electronics to provide a measurement of the centroid of the pulse height distribution and plan to test the prototype and associated electronics in the next few months.

As part of our proposal, we request that the Laboratory provide the following support:

1) An additional coil for our analyzing magnet in order to increase the present magnetic field from 11 kG to 18 kG.

2) An upgrading of the present beam to P-East which would allow us to have beam intensities as high as 2×10^{13} protons on our target. We have not yet been able to use the full intensity of the beam. The beam losses upstream of the target box are a source of muons which ultimately limit the intensity at which we run. We are aware of the effort to reduce these losses by increasing the aperture

of the quadrupoles in P-East. We are hopeful that this will eliminate the difficulty. If it does not, a more drastic solution such as the one which has been carried out in P-West may be necessary.

3) We also request that the laboratory provide support for the construction of the polarizer. We would actively participate in the design, construction and testing of the device. The cost of the graphite crystal we need will be of the order of $ 20 K.

4) We also request that the laboratory provide a cryogenic target 2 1/2 in. diameter, 24 in. long.

IV. REQUEST FOR RUNNING TIME

We divide our request for running time into four categories.

1. A Study of the Reactions:

$$\gamma + \mathrm{Be} \rightarrow \alpha^+ + \alpha^-$$
$$\hookrightarrow \pi^+\pi^+\pi^-\pi^0\pi^0 + \ldots$$
$$\hookrightarrow K^+\pi^-\pi^-$$
$$\hookrightarrow \mu\pi^+\nu,\ e\pi^+\nu \quad \text{etc., etc.}$$

$$\gamma + \mathrm{Be} \rightarrow 1^- + \mathrm{Be}$$
$$\hookrightarrow \pi^+\pi^-$$
$$\hookrightarrow \pi^+\pi^+\pi^-\pi^-$$
$$\hookrightarrow \pi^+\pi^+\pi^-\pi^-\pi^0 \quad \text{etc., etc.}$$

For 400 hours of data taking, we should be able to collect $\sim 10^6$ events for $E_\gamma \geqslant 100$ GeV.

2. A Study of the Reaction:

$$\gamma + p \rightarrow \mu^+\mu^- + \text{anything} .$$

We are interested in particular in the measurement of the slope of the differential cross section for the $\psi(3100)$. We are also interested in the final states, such as $\mu^+K^-\mu^-$---etc. For 400 hours, we should be able to collect 1000 $\psi \rightarrow \mu^+\mu^-$.

3. An initial test of the linearly polarized photon beam.

The essential aim of this part of the proposal is to install and test the polarizer in the wide band neutral beam. We will concentrate on measurements of the 1^- states. In particular, we will measure the production of the ρ^o and other 1^- states: 400 hours.

4. One interesting aspect of our experiment has been the ability to measure simultaneously the hadronic production of dimuons along with the photoproduction. We propose to continue our study of the continuum along with further studies of $\psi(3100)$ and $\psi'(3700)$. We should obtain $\sim$ 500 $\psi \rightarrow \mu^+\mu^-$ and $\sim$ 10 $\psi' \rightarrow \mu^+\mu^-$: 300 hours.

<u>Total hours: 1500 hours.</u>

We would like to take data by Spring of 1976.

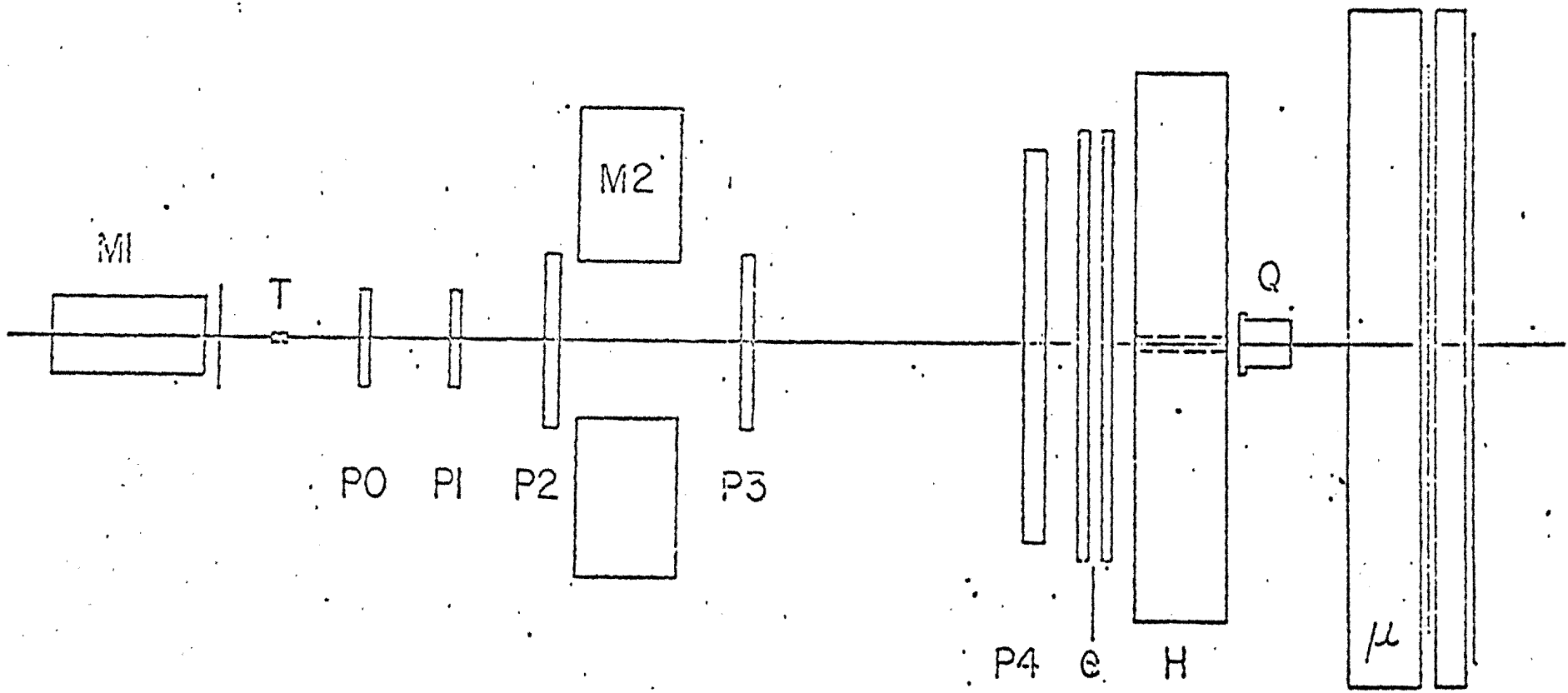


Fig 1

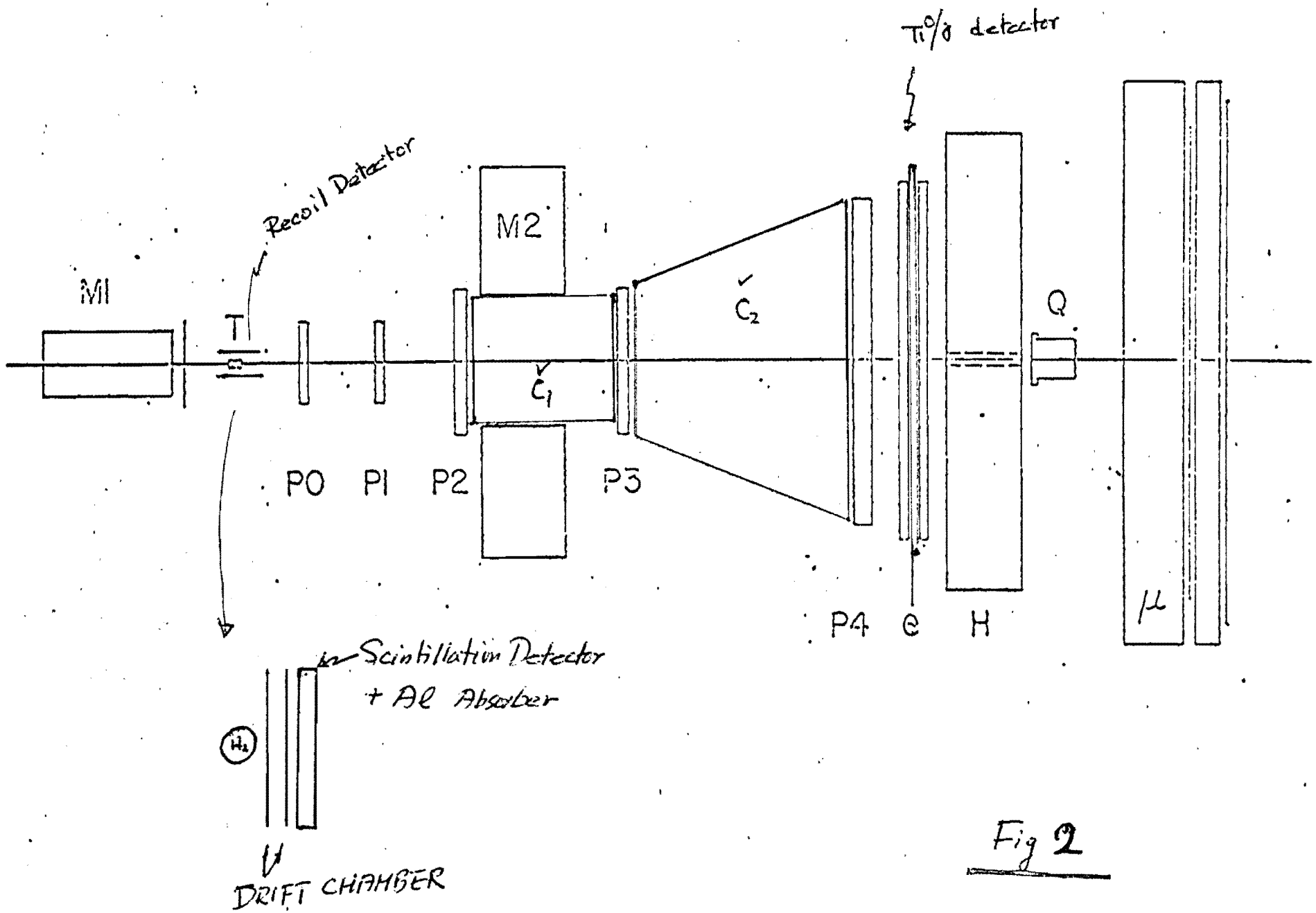

Ti°/γ detector
Recoil Detector
M1
T
M2
C1
C2
Q
P0
P1
P2
P3
P4
e
H
μ
Scintillation Detector
+ Al Absorber
H2
DRIFT CHAMBER

Fig 2

APPENDIX I

The availability of intense beams of photons with energies of tens and hundreds of GeV increases the practicality of linearly polarizing photons by transmission through anisotropic absorbers. Physicists have long predicted interference phenomena that would modify rates of electron-positron pair production by high energy photons in crystalline targets.[1-4] The total rate of pair production, and thereby photon attenuation, can depend on the photon linear polarization direction, a result of enhancement of the pair production rate for special values of target recoil momentum and a correlation between the directions of photon linear polarization and the recoil momentum. Surprisingly enough, the maximum achievable interference contributions to the attenuation rate increase linearly with photon energy for an ideal crystal.

Two experiments have confirmed the polarization predictions for photons of 8.7 GeV[5] and 16 GeV.[6] The latter experiment achieved 25% linear polarization with a factor of about 20 reduction in the intensity of 16 GeV photons. Such polarization and energy dependent attenuation leads to predictions[7] of birefringence properties that could conceivably be used to circularly polarize photons. No experiments to demonstrate the feasibility of such circular polarization have been attempted. Predictions of crystal interference for electron bremsstrahlung, a process related to pair production and covering the same range in target

recoil momentum, have been verified in great detail. Thin crystal targets have been used in roughly one to ten GeV electron beams to provide highly polarized mono-energetic photons, typically with less than half the initial electron energy.

We wish to review the existing calculations and observations to determine possible use of these interference phenomena in a high energy photon beam now in use at the National Accelerator Laboratory.[8] Relevant equations and derivations are scattered throughout twenty years of physics publications. We shall attempt only to summarize the results in a manner suitable for numerical calculations of observable effects and for qualitative understanding of the many underlying physical phenomena and the approximations used in their analysis. Except as noted, we shall use units in which $\hbar$, c and electron mass m_e are all equal to 1.

I. Pair Production in Crystals

Except for the complication introduced by the two-lepton final state amplitudes, the description of coherent pair production from crystals is quite similar to descriptions of x-ray crystallography. For a single crystal of N identical atoms in a volume V, the effects of crystal structure, in the first Born approximation, can be isolated in a single factor

$$I(\vec{q}) = \frac{1}{V} \left| \sum_{i=1}^{N} e^{i\vec{q}\cdot\vec{x}_i} \right|^2 \tag{1}$$

where $\vec{q}$ is the momentum transferred to the target and the sum is over the coordinates $\vec{x}_i$ of all N atoms. The differential cross section per unit volume can then be written as the differential cross section per atom multiplied by $I(\vec{q})$:

$$\frac{\partial^3 \mu}{\partial \vec{q}} = I(\vec{q}) \frac{\partial^3 \sigma}{\partial \vec{q}} . \tag{2}$$

The volume cross section μ is simply the fraction of photons per unit length converted to electron-positron pairs. The atomic cross section σ should be the cross section per atom for producing an electron-positron pair. In fact, we shall consider only the coherent single atom process in which the amplitudes for scattering from the nucleus and the atomic electrons add coherently and the atom remains in its ground state. For light atoms, we may be neglecting noticable contributions from inelastic processes in which a valence electron is excited from one crystal state to another. The factor I(q) appears in the same manner in x-ray scattering from crystals of identical atoms.

For a quick intuitive picture of the interference effects, consider a one-dimensional scattering problem. If N scattering centers are equally spaced by a distance a for a total crystal length L=Na, then

$$I(q) = \left(\frac{1}{L}\right) \left| \sum_{n=1}^{N} e^{ignna} \right|^2 = \left(\frac{1}{L}\right) \left(\frac{\sin Nq\frac{a}{2}}{\sin q\frac{a}{2}}\right)^2 \tag{3}$$

$$\lim_{N\to\infty} I(q) = 2\pi \left(\frac{N}{L}\right)^2 \sum_{n=0}^{\infty} \delta\left(q-2\pi\frac{n}{a}\right) \tag{4}$$

For large N, I(q) is negligible except for sharp peaks about discrete values of q. Each peak has area $2\pi(N/L)^2$, height N^2/L, and width $2\pi/L$. Over any range of q large compared to the discrete spacing $2\pi/a$, the average value of I(q) is N/L, so that if the differential cross section $\frac{d\sigma}{dq}$ for scattering by a single point varies sufficiently slowly with q then the total volume cross section is $\mu=(N/L)\sigma$. If on the other hand, $\frac{d\sigma}{dq}$ is a sharply peaked function, say of width t about $q=q_o$, then μ depends strongly on a if t is less than the spacing $2\pi/a$. If $q_o=2\pi n/a$ for some interger n, then μ can be simply estimated in the two limits

$$\frac{Lt}{2\pi} >> 1, \qquad \mu \sim 2\pi\left(\frac{N}{L}\right)^2\left(\frac{d\sigma}{dq}\right)_{q=q_o} \tag{5}$$

$$\frac{2\pi}{Lt} >> 1, \qquad \mu \sim tL\left(\frac{N}{L}\right)^2\left(\frac{d\sigma}{dq}\right)_{q=q_o} \tag{6}$$

We have discovered a <u>coherence length</u>, $L_c=2\pi/t$. If our crystal is shorter than this length, the number of interactions per unit length depends on the length of the crystal.

Suppose next that these scattering centers have uncorrelated motion about their fixed average positions. Consider the coordinates of the individual scattering centers to be $x_n=na+u_n$, where u_n is a random variable with a probability distribution $P(u_n)=1/\sqrt{2\pi A}\, e^{-u_n^2/2A}$, so that $\langle u_n^2\rangle = \int u_n^2 P(u_n)\,du_n$ is simply A. I(q) then becomes

$$I(q) = \frac{1}{L}\int P(u_1)du_1 \ldots \int P(u_n)du_n \left| \sum_{n=1}^{N} e^{iq(na+u_n)} \right|^2$$

$$= \left(\frac{N}{L}\right) + \left(\frac{1}{L}\right) \sum_{m \neq n} \int P(u_m)du_m \int P(u_n)du_n e^{iq(na+u_n-ma-u_m)}$$

$$= \left(\frac{N}{L}\right) + \left(\frac{1}{L}\right) e^{-Aq^2} \left(\left| \sum_{n=1}^{N} e^{iqna} \right|^2 - N \right) \tag{7}$$

In the limit that A is large compared to a^2, no periodic structure remains and we have the expected result that $I(q)=(N/L)$.

Extension to three dimensions is straightforward. An ideal crystal can be considered an array of identical unit cells. Three independent vectors $\vec{a}_i$ form a basis of the <u>direct lattice</u>. Any point $\vec{x}$ in the lattice is the sum of integral multiples of the basis vectors $\vec{a}_i$. We then define three vectors $\vec{b}_i$, so chosen that $\vec{a}_i \cdot \vec{b}_j$ is 2π if i=j, and is zero otherwise:

$$b_i = 2\pi \frac{\vec{a}_j x \vec{a}_k}{\vec{a}_i \cdot (\vec{a}_j x \vec{a}_k)} \tag{8}$$

where ijk is a cyclic permutation of 123. The vectors $\vec{b}_i$ form a basis of the <u>reciprocal lattice</u>. Any vector $\vec{g}$ in the reciprocal lattice is the sum of integral multiples of the b_i:

$$g_{hk\ell} = h\vec{b}_1 + k\vec{b}_2 + \ell\vec{b}_3 \tag{9}$$

where h, k, and ℓ are integers. If $\vec{x}$ is any point in the direct lattice and $\vec{g}$ any vector in the reciprocal lattice,

then $\vec{g}\cdot\vec{x}$ is an integral multiple of 2π and e^{igx} is just unity. We also define a structure factor $S(\vec{g})$:

$$S(\vec{g}) = \frac{1}{N_o} \sum_{i=1}^{N_o} e^{i\vec{g}\cdot\vec{x}_i} \tag{10}$$

where the summation is over the coordinates $\vec{x}_i$ of all N_o atoms in the unit cell.

An approximation for I(q) can now be presented:

$$I(q) \approx n - ne^{-\underset{\sim}{A}\vec{q}\vec{q}} + (2\pi)^3 n^2 e^{-\underset{\sim}{A}\vec{q}\vec{q}} |S(\vec{q})|^2 \sum_{hk\ell} \delta(\vec{q}-g_{hk\ell}) \tag{11}$$

where n is the number of atoms per unit volume, and the final sum of Dirac delta functions is over all momenta $\vec{g}_{hk\ell}$ of the crystal reciprocal lattice. The Debye-Waller factor $e^{-\underset{\sim}{A}\vec{q}\vec{q}}$ is an approximation for the effects of lattice vibrations, and $\underset{\sim}{A}\vec{q}\vec{q}$ is meant to indicate the bilinear sum $\sum_{ij} A_{ij}q_iq_j$. For beryllium and graphite, and for any crystal with exactly one axis of more than two-fold symmetry, the most general form of $\underset{\sim}{A}\vec{q}\vec{q}$ is $Aq_{\parallel}^2 + A'q_{\perp}^2$, where $q_{\parallel}$ and $\vec{q}_{\perp}$ are components of $\vec{q}$ respectively parallel and perpendicular to this symmetry axis. For crystals with more than one such symmetry axis, $\underset{\sim}{A}\vec{q}\vec{q}$ reduces to simply Aq^2.

The volume cross section can be written as the sum of three terms:

$$\mu = \mu_0 - \mu' + \mu_I. \tag{12}$$

The first term, $\mu_0 = n\sigma$, is the volume cross section for an amorphous material, where the contributions of individual atoms are added incoherently. The second term

$$\mu' = n \int d^3q e^{-\underset{\sim}{A}\vec{q}\vec{q}} \frac{\partial^3\sigma}{\partial\vec{q}} \tag{13}$$

is an interference term that reduces the incoherent contributions. If $\underset{\sim}{A}\vec{q}\vec{q}$ is simply Aq^2, μ' is independent of crystal orientation for fixed photon momentum $\vec{k}$ and polarization direction $\hat{\epsilon}$. The third term,

$$\mu_I = (2\pi)^3 n^2 \sum_{hk\ell} |S(\vec{g})|^2 e^{-\underset{\sim}{A}\vec{g}\vec{g}} \frac{\partial^3\sigma}{\partial\vec{q}}\Bigg)_{\vec{q}=\vec{g}} \tag{14}$$

depends explicitly on crystal orientation and may therefore depend on photon linear polarization through polarization-dependence of the atomic differential cross section. In fact, the atomic cross section favors momentum transfer $\vec{q}$ nearly perpendicular to the photon momentum $\vec{k}$ and perpendicular to the photon electric field $\vec{\epsilon}$, permitting crystal orientations relative to the photon momentum $\vec{k}$ for which a single momentum $\vec{g}$ in the reciprocal lattice dominates the expression for μ_I. The crystal then attenuates most rapidly photons linearly polarized perpendicular to $\vec{g}$.

Details of crystal structure sufficient for numerical calculations are described in Appendix A.

To proceed further, we must examine the atomic differential cross sections. Ref. 2 outlines numerical calculations of polarization-dependent total pair production rates in crystals. Ref. 3 concentrates on electron bremsstrahlung in crystals but outlines calculations of

$$\frac{d\mu}{dy} = \int I(\vec{q})\frac{\partial^4\sigma}{\partial y\partial\vec{q}}\, dq^3 \tag{15}$$

where y is the fraction of the initial photon energy carried away by a single lepton. This latter approach is particularly useful if one wishes to detect the lepton pairs, but the former is an easier method for calculation of photon attenuation. From Ref. 3, we have extracted the differential cross section for elastic pair production from a single neutral atom of atomic number Z, calculated in the Born, high energy and low q^2 approximations:

$$\frac{\partial^4\sigma}{\partial y\partial^3\vec{q}} = \left(\frac{\sigma_o k}{2\pi}\right) f(q^2)\left(\frac{q_\perp}{q}\right)^2\left(\frac{2}{k\cdot q}\right)^2\left[k\delta-\left(1-\frac{\delta}{q_{||}}\right)^2-2(\hat{\epsilon}\cdot\hat{q}_\perp)^2\left(\frac{\delta}{q_{||}}\right)^2\right]$$
$$\theta\left(q_{||}-\delta-\frac{q^2}{2k}\right) \tag{16}$$

where $\sigma_o = Z^2\left(\frac{e^2}{\hbar c}\right)^3\left(\frac{\hbar}{m_e c}\right)^2 = (0.5794 \times 10^{-27}\text{cm}^2)Z^2$

$\vec{k}$ = photon momentum

$\vec{q}_\perp$, $q_{||}$ = components of momentum transfer $\vec{q}$ perpendicular, parallel to $\vec{k}$

$\delta = 1/[2ky(1-y)]$ is the kinematic lower limit for $q_{||}$

$\theta(x) = 1$ if $x>0$

$\quad\;\; =0$ if $x<0$

$\hat{\epsilon}$ is the photon electric field direction, so that $\hat{\epsilon}\cdot\hat{q}$ is the cos ϕ, where ϕ is the angle between the photon linear polarization vector and $\vec{q}_\perp$.

$f(q^2) = (1-F(q^2))^2/q^2$ where

$F(q^2)$ is the atomic electron form factor for the free neutral atom.

For our purposes, we can ignore q^2 compared to 1 or $k\cdot q$ and consider $q_\perp^2$ equal to q^2. We then write the cross section in the form

$$\frac{\partial^4\sigma}{\partial y\partial^3 q} = \left(\frac{k\sigma_o}{2\pi}\right) f(q^2)[h_1(x,u)-h_2(x,u)\cos 2\phi] \tag{17}$$

where $h_1(x,u) = x^2[2u-(1-xu)^2-(xu)^2]$

$h_2(x,u) = x^2[(xu)^2]$

$x = 2/k\cdot q$

$u = 1/[4y(1-y)]$

with the restrictions $0<x\leq xu\leq 1$. Note that the polarization dependence is greatest for $x\approx 1$ which requires symmetric electron pairs, $y\approx 1/2$. We shall see later that for real crystals and very large photon energies, calculations of μ_I require integration of h_1 and h_2 over $\vec{k}\cdot\vec{q}$, resulting in the y-dependence :

$$\frac{d\mu_I}{dY} \propto 1-\left(\frac{2}{3}\right)y(1-y)(2+\cos 2\phi) \tag{18}$$

We next integrate the cross section over y, to obtain, in agreement with the low q^2 limit of the more detailed formula of Reference 2:

$$\frac{\partial^3\sigma}{\partial\vec{q}} = \left(\frac{k\sigma_o}{2\pi}\right) f(q^2)[h_1(x)-h_2(x)\cos 2\phi] \tag{19}$$

where $h_1(x) = x^2[(1+x-\frac{x^2}{2})R-(1+x)\sqrt{1-x}\,]\theta(\frac{1}{x}-1)$

$h_2(x) = (\frac{x^3}{2})[\sqrt{1-x} + (\frac{x}{2})R]\theta(\frac{1}{x}-1)$

$R = \ln[(1+\sqrt{1-x})/(1-\sqrt{1-x})]$

The coherent pair production cross section per unit volume then becomes

$$\mu_I = k \sum_{\vec{g}} F(\vec{g})[h_1(2/\vec{k}\cdot\vec{g})+(1-2(\hat{\varepsilon}\cdot\hat{g})^2)h_2(2/\vec{k}\cdot\vec{g})] \tag{20a}$$

where $F(\vec{g}) = (2\pi)^2\sigma_o n^2|S(\vec{g})|^2 e^{-A\vec{g}\vec{g}} f(g^2)$. (20b)

If we choose a coordinate system with the z axis parallel to $\vec{k}$, the spatial development of a photon amplitude u(z) obeys

$$\frac{du}{dz} = i(k+\Delta+i\mu/2)u. \tag{21}$$

The real and imaginary parts of the crystal index of refraction are, respectively, $1+\Delta/k$ and $\mu/2k$. The Kramers-Kronig dispersion relations then require

$$\Delta(k,\hat{\varepsilon}) = \frac{k}{\pi} P\int_0^\infty \frac{\mu(\kappa,\hat{\varepsilon})d\kappa}{\kappa^2-k^2} \tag{22}$$

where P denotes Cauchy principle value.

If we ignore all processes, except coherent pair production, that could reduce the intensity of photons with momentum $\vec{k}$, we find

$$\Delta(k,\hat{\varepsilon}) = k \sum_{\vec{g}} F(\vec{g})[d_1(2/\vec{k}\cdot\vec{g})+(1-2(\hat{\varepsilon}\cdot\hat{g}_\perp)^2)d_2(2/\vec{k}\cdot\vec{g})] \tag{23}$$

where
$$d_i(2/kg_{||}) = \frac{1}{\pi}P_o\int^\infty \frac{\kappa h_i(2/\kappa g_{||})d\kappa}{\kappa^2-k^2}$$
$$= \frac{1}{\pi}\left(\frac{2}{kg_{||}}\right)^2 P_o\int^1 \frac{h_i(x)dx}{x((2/kg_{||})^2-x^2)}. \tag{24}$$

Since we are only interested in polarization dependence of the phase shifts, we evaluate only the integral for d_2, obtaining, as in Ref. 3:

$$d_2(x) = \frac{x^2}{2\pi}(A+B-C)\theta(x)$$

$$A = \{\sqrt{x-1} - \cot^{-1}\sqrt{x-1}\}^2\theta(x-1)$$

$$B = [\{\frac{\pi x}{2}\}^2 - \{\sqrt{1-x} + (\frac{x}{2})\ln(\frac{1+\sqrt{1-x}}{1-\sqrt{1-x}})\}^2]\theta(1-x)$$

$$C = \{\sqrt{1+x} - (x/2)\ln(\frac{\sqrt{1+x}+1}{\sqrt{1+x}-1})\}^2$$

For specified crystal orientation and photon momentum $\vec{k}$, we may write,

$$\mu_I = \mu_1 + (1-2(\hat{\epsilon}\cdot\hat{t})^2)\mu_2 \tag{26a}$$

$$\Delta = \Delta_1 + (1-2(\hat{\epsilon}\cdot\hat{t}')^2)\Delta_2. \tag{26b}$$

Fairly straightforward numerical evaluation of the previous expressions yields μ_1, μ_2, Δ_2 and the directions $\hat{t}$ and $\hat{t}'$. In general, $\hat{t}$ and $\hat{t}'$ may depend on photon momentum and may not be simply related to each other.

The functions h_1, h_2 and d_2 are plotted in Figure 1. Figure 2 presents $f(q^2)$ for various elements, using atomic electron form factors from tables based on Hartree-Fock calculations for neutral free atoms.[9] The details of the crystal electron charge distribution are important for momentum transfer comparable to inverse interatomic distances. Since the spatial distribution of the valence electrons in a crystal is seldom spherically symmetric, one should expect noticable departures from the spherical symmetry of the free atomic form factors.

To examine the dependence of μ and Δ on crystal orientation, first note that the greatest crystal plane spacings

are usually one to three Angstroms, resulting in reciprocal lattice spacings on the order of .01 m_e. On this scale, the contributions of a specific lattice point $\vec{g}$ are very sharp functions of $g_{||}$, with a coherence length along the beam direction of order k/m_e^2 or roughly 1000 Angstroms for 100 GeV photons. Significant contributions are limited by the relatively slowly varying dependence on $g_\perp$ to a region $g_\perp \leq .05\ m_e$. It is often useful to consider interference effects noticable only for momentum transfer in the very thin disc $q_\perp < .05\ m_e$ and $0 < q_{||} < 10(\frac{m_e}{k})m_e$. For most crystal orientations, no reciprocal lattice vectors fall within the disc, and μ_I and Δ_2 are negligible.

If the photon momentum $\vec{k}$ is very nearly parallel to a single family of parallel crystal planes, with plane spacing d, then the disc of important momentum transfer grazes a line of reciprocal lattice vectors spaced by $2\pi/d$ along the normal to the planes. This normal is then the single axis of anisotropy for both μ and Δ. Figure 3 shows the variation of μ_I and Δ_2 with the angle θ between $\vec{k}$ and the (001) planes, the family of planes perpendicular to the c-axis in beryllium and in graphite, for 50 GeV photons and room temperature crystals. We plot $\mu_\perp$ and $\mu_{||}$, the values of μ_I for photon linear polarization respectively perpendicular and parallel to the c-axis. For comparison, we have also indicated the value of μ_o, the expected attenuation constant for amorphous material.[10] Note that the previous expressions for μ_I and Δ_2 supply a simple scaling law, that the magnitudes

increase linearly with photon energy while the angular scale is inversely proportional to photon energy. With this scaling rule, we can use Figure 3 for arbitrary k and θ, so long as other families of crystal planes can be ignored.

The maximum interference effects attainable with real crystals are limited by a crystal imperfection called the <u>mosaic spread of crystal axes</u>. A macroscopic crystal can often be considered a collection of smaller crystals with some variation in the orientation of their well-defined axes. Consider a crystal axis with a gaussian distribution

$$W(\theta) = \frac{1}{\theta_o\sqrt{2\pi}} e^{-\frac{1}{2}(\theta/\theta_o)^2} \tag{27}$$

for the angle of rotation θ of the smaller crystals about another fixed axis orthogonal to the first. The fraction of the smaller crystals rotated by an angle between θ and θ $\theta+d\theta$ is then $W(\theta)d\theta$.

The calculations of μ_I and Δ must therefore be averaged over the angular distribution of the locally defined crystal axes:

$$\overline{\mu_I}\ (\bar{\theta}) = \int_{-\infty}^{\infty} W(\theta-\bar{\theta})\mu_I(\theta)d\theta \tag{28}$$

The graphite used in Ref. 5 and 6, for example, is probably the best obtainable and has c-axis mosaic spreads described by θ_o of 3 to 4 milliradians.[11] For beryllium θ_o can probably be pushed well below a milliradian. A

lower limit for desired mosaic spread may be imposed by mechanical alignment tolerances, by photon beam divergence, and possibly by the photon momentum spectrum.

The interference effects in such crystals should become independent of photon energy for sufficiently large energy. We can estimate for instance, for $kg\theta_o \gg 1$,

$$\bar{\mu}_i(\bar{\theta}) = \int_{-\infty}^{\infty} W(\bar{\theta}-\theta)\{k\sum_{n=1}^{\infty} F(ng)h_i(2/ngk|\theta|)\}d\theta \tag{29a}$$

$$\approx W(\bar{\theta})\{\sum_{n=1}^{\infty}(\frac{2}{ng})F(ng)\}2\int_0^{\infty} h_i(\frac{1}{y})dy. \tag{29b}$$

The integrals are easily evaluated, yielding

$$\bar{\mu}_1(\bar{\theta}) \approx (28/9)W(\bar{\theta})\{\sum_{n=1}^{\infty}(\frac{2}{ng})F(ng)\} \tag{30a}$$

$$\bar{\mu}_2(\bar{\theta}) \approx (1/7)\bar{\mu}_1(\bar{\theta}). \tag{30b}$$

Figure 4 shows the momentum dependence of $\bar{\mu}_1$ $\bar{\mu}_2$ and $\bar{\Delta}_2$ for room temperature graphite with mosaic spread θ_o = 3.5 mr for some small values of $\bar{\theta}$, i.e., c-axis perpendicular to $\vec{k}$. Figure 5 shows the same curves for room temperature beryllium with θ_o = 1.0 mr. These curves can be scaled for arbitrary mosaic spread with the observation that $\bar{\mu}_1$, $\bar{\mu}_2$ and $\bar{\Delta}_2$ are inversely proportional to θ_o, for fixed $k\theta_o$ and $\bar{\theta}/\theta_o$.

We again emphasize that these first examples consider only the contributions from a single family of parallel crystal planes. If k is also very nearly parallel to a

second family of parallel crystal planes, the disc of important moment transfer grazes a two dimensional grid of reciprocal lattice vectors in the plane containing the normals to each of the families of planes. In this case, μ_I can become much larger, but, since the favored momentum transfer no longer has fixed direction, the anisotropy may even decrease. In general, $\hat{t}$ and $\hat{t}'$ vary with energy.

Figure 6 shows the energy dependence of $\bar{\mu}_1$, $\bar{\mu}_2$ and $\bar{\Delta}_2$ for a room temperature beryllium crystal with gaussian mosaic spread θ_o = 1 mr root-mean-square rotation about any axis, and so oriented with respect to a photon beam that $\vec{k}$ forms a mean angle θ_1 with the (001) planes and a mean angle θ_2 with the (100) planes.

II. Photon Polarization

For photons of momentum $\vec{k} = kz$, with x,y components of transverse electric field, we first factor out an overall phase shift and attenuation from the two component amplitude

$$\begin{pmatrix} \epsilon_x(z) \\ \epsilon_y(z) \end{pmatrix} = u(z)\ e^{i(k-\mu/2)z} \quad . \quad e^{(ik-\mu/2)Z} \ . \tag{31}$$

Recall the Pauli spinors

$$\sigma_o = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{32}$$

and note that a rotation of the coordinate system by an angle ϕ about the z axis is accomplished with the operator

$$R(\phi) = e^{i\phi\sigma_2} \ . \tag{33}$$

If the axes of anisotropy for absorption and phase shift have been rotated away from the x axis by ϕ and ϕ'

respectively,

$$\frac{du}{dz} = [i\Delta_2 R(-\phi')\sigma_3 R(\phi') - (\frac{\mu_2}{2}) R(\phi)\sigma_3 R(-\phi)]u(z) \quad . \tag{34}$$

Henceforth, we consider only

$$\begin{aligned} &\phi = \phi' = \phi_o + \lambda z \\ &\phi_o, \lambda, \mu_2, \Delta_2 \quad \text{real and constant} \\ &n = \Delta_2 + i\mu_2/2 \quad . \end{aligned} \tag{35}$$

So the propagation equation becomes

$$\frac{du}{dz} = inR(-\phi)\sigma_3 R(\phi)u(z) \quad . \tag{36}$$

For discussion of partial polarization, the incoherent super position of different amplitudes $u_n(z)$, we introduce the density matrix

$$\rho(z) = \sum_n u_n(z) u_n(z)^\dagger \tag{37a}$$

$$= A_o(z)\sigma_o + \vec{A}(z)\cdot\sigma \quad . \tag{37b}$$

The four real numbers A_α are the stokes parameters. The photon intensity is $I(z) = e^{-\mu z}A_o(z')$ and $\vec{A}(z)/A_o(z)$ fully describes the polarization. A definite single state is fully polarized, with $|\vec{A}| = A_o$, but a partially polarized beam has $|\vec{A}| < A_o$. The circular polarization is A_2/A_o, and A_1/A_o, A_2/A_o describe linear polarization.

To study the propagation equation (36), we choose to rotate the coordinate system to keep the axis of anisotropy of the polarizing medium always along the x-axis:

$$\bar{u}(z) = R(\phi)u(z) \tag{38}$$

$$\bar{\rho}(z) = R(\phi)\rho(z)R(-\phi) \tag{39a}$$

$$= A_o(z) + \bar{A}(z)\cdot\sigma \ . \tag{39b}$$

(Note that A_o and A_2 are invariant under R.)

Since we have allowed z-dependence for ϕ,

$$\frac{d\bar{u}}{dz} = i(\lambda\sigma_2+n\sigma_3)\bar{u}(z) \tag{40}$$

with solution

$$\bar{u}(z) = \bar{M}(z)\bar{u}(o) \tag{41}$$

where $\bar{M}(z) = \sigma_o\cos\kappa z + \frac{i}{k}(\lambda\sigma_2+n\sigma_3)\sin\kappa z$ (42)

and the complex constant κ satisfies

$$\kappa^2 = \lambda^2 + n^2 \ . \tag{43}$$

The density matrix in the rotated system

$$\bar{\rho}(z) = A_o(z)\sigma_o + \vec{\bar{A}}\cdot\vec{\sigma} \tag{44}$$

is therefore

$$\bar{\rho}(z) = \bar{M}(z)\bar{\rho}(o)\bar{M}(z)^{\dagger} \ . \tag{45}$$

We can write explicitly a differential equation for $\bar{\rho}$,

$$\frac{d}{dz}\bar{\rho}(z) = i[\lambda\sigma_2+n\sigma_3]\bar{\rho}(z)-i\bar{\rho}(z)[\lambda\sigma_2+n^*\sigma_3] \tag{46}$$

or

$$\frac{d}{dz}\begin{pmatrix} A_o \\ \bar{A}_1 \\ A_2 \\ \bar{A}_3 \end{pmatrix} = \begin{bmatrix} 0 & 0 & 0 & -\mu_2 \\ 0 & 0 & 2\Delta_2 & -2\lambda \\ 0 & 2\Delta_2 & 0 & 0 \\ -\mu_2 & 2\lambda & 0 & 0 \end{bmatrix} \begin{bmatrix} A_o \\ \bar{A}_1 \\ A_2 \\ \bar{A}_3 \end{bmatrix} \tag{47}$$

We can see that Δ_2 and λ rotate $\vec{\bar{A}}$ about its 3-axis and 2-axis respectively, and that only μ_2 affects A_o or the

magnitude of $\bar{A}$. Net linear polarization along the axis of anisotropy is achieved by μ_2, but Δ_2 generates circular polarization from linear polarization at 45° to the axis of anisotropy. Discrete or continuous changes in the axis of anisotropy are therefore necessary to achieve circular polarization. (Different axes of anisotropy for phase shift and attenuation could suffice, but for crystal polarizers this seems to require impractically low values of μ_2 and Δ_2.)

A practical polarizer would consist of several short crystals, and so we consider successive transmissions through crystals of length L_j and fixed orientations ϕ_j, for which reduces to

$$\bar{M}_j = \sigma_o \cos(nL_j) + i\sigma_3 \sin(nL_j) \tag{48}$$

or, in a fixed coordinate system,

$$M_j = R(-\phi_j)\bar{M}_j R(\phi_j) . \tag{49}$$

Transmission through a series of N such crystals of total length L is then described by

$$M(L) = \prod_{j=1}^{N} M_j \tag{50}$$

$$\rho(L) = M(L)\rho(o)M(L)^\dagger . \tag{51}$$

For an initially unpolarized beam transmitted by a uniform absorber with axis of anisotropy fixed along the x-axis, we obtain the familiar results for intensity I(z) and linear polarization P(z):

$$I(z) = I_o e^{-\mu z} \cosh(\mu_2 z) \tag{52a}$$

$$P(z) = \tanh(\mu_2 z) \ . \tag{52b}$$

For a numerical example, consider graphite in Fig. with $\bar{\theta} = \theta_o$. For k between 50 and 250 GeV, μ_2 varies from 0.94/m to 1.17/m, and $\mu = \mu_o + \mu_1$ varies from 7.6/m to 10.7/m. For 25 cm of absorber, we would expect about 25% linear polarization for all photons above 50 GeV, with a attenuation varying from a factor of 6.7 at 50 GeV to a factor of 15 at 250 GeV, with attenuation by only 2.8 for low energy photons. Thus, graphite crystals similar to ones already in use [6] could be positioned with readily achievable tolerances to provide useful linear polarization simultaneously for all photons above some lower limit. Increasing the low energy cutoff would permit less attenuation for given polarization.

Circular polarization is more difficult to describe or implement. R. Friedberg has determined the variation of $\phi(z)$ required to maximize $A_2(L)$ for fixed μ_2, Δ_2 and L.[11] He discusses practical considerations for circular polarization and concludes that, for $|n|L << 1$, the best distribution is

$$\phi(z) = (\frac{\pi}{2L})z \qquad |n|L << 1 \ . \tag{53}$$

With the result, for $\kappa = \alpha + i\beta \approx \lambda + i\,\frac{\mu_2\Delta_2}{2\lambda}$, (54)

$$A_o(L) = |\cos\kappa L|^2 + (\frac{\lambda^2 + |n|^2}{|\kappa|^2})|\sin\kappa L|^2 \approx 1 \ . \tag{55}$$

$$A_2(L) = \left(\frac{\lambda}{\alpha^2+\beta^2}\right)[\beta\sin^2\alpha-\alpha\sinh 2\beta] \approx -\left(\frac{2}{\pi}\right)\mu_2\Delta_2 L^2 \quad . \tag{56}$$

For the graphite considered above, Δ_2 is 0.5/m at 50 GeV, so that 25 cm could yield almost 2% circular polarization at 50 GeV and considerably less at 250 GeV. Circular polarization is a second-order effect and must always be less than the possible linear polarization for fixed attenuation in similar crystals. In addition, because Δ_2 must go to zero for infinite photon energy, only a finite range photon energy is simultaneously circularly polarizable.

Truly practical circular polarization at energies above 100 GeV could be achieved with crystals of smaller mosaic spread and correspondingly tighter alignment tolerances. The first sections of such a polarizer should maximize linear polarization, while the last sections should have $\bar{\theta} = 0$ to better accuracy than θ_o, and θ_o as small as possible. (Consider Fig. 5 and recall that the horizontal and vertical scales are inversely proportional to θ_o.)

APPENDIX A

Physical Parameters of Crystals

A1. Specific Representation of Lattice Bases

We consider only cubic or hexagonal crystal structures and provide specific representations of the lattice vectors in a cartesian coordinate system. By $\hat{x}_i$, we designate a unit vector along one of three mutually orthogonal axes. For cubic structures, the direct and reciprocal lattice vectors can be written

$$\vec{a}_i = a\hat{x}_i$$

$$\vec{b}_i = \left(\frac{2\pi}{a}\right)\hat{x}_i$$

For hexagonal structures, we use the cartesian representation

$$\vec{a}_1 = \left(\frac{a}{2}\right)(\sqrt{3}\,\hat{x}_1 - \hat{x}_2) \qquad \vec{b}_1 = \left(\frac{2\pi}{a}\right)\left(\frac{2}{\sqrt{3}}\right)\hat{x}_1$$

$$\vec{a}_2 = a\hat{x}_2 \qquad \vec{b}_2 = \left(\frac{2\pi}{a}\right)(\hat{x}_2 + \hat{x}_1/\sqrt{3})$$

$$\vec{a}_3 = c\hat{x}_3 \qquad \vec{b}_3 = \left(\frac{2\pi}{c}\right)\hat{x}_3$$

A2. Unit Cell and Structure Factors

Designate an atom's position within a unit cell, $\vec{x} = r\vec{a}_1 + s\vec{a}_2 + t\vec{a}_3$, by (rst). Consider a reciprocal lattice vector $g_{hk\ell} = h\vec{b}_1 + k\vec{b}_2 + \ell\vec{b}_3$. The structure factor is then

$$S = \frac{1}{N_o} \sum_{(rst)} e^{i2\pi(rh+sk+t\ell)}$$

Body centered Cubic Structure (bcc)

The bcc structure has two atoms per unit cell, (000)

and (1/2 1/2 1/2), and so the structure factor is

$$S(hk\ell) = \frac{1}{2}(1+e^{i\pi(h+k+\ell)})$$

$$= 1 \text{ if } h+k+\ell \text{ is even}$$

$$= 0 \text{ otherwise}$$

Face Centered Cubic Structure (fcc)

The fcc structure has four atoms per unit cell, (000), $(0\frac{1}{2}\frac{1}{2})$, $(\frac{1}{2}0\frac{1}{2})$ and $(\frac{1}{2}\frac{1}{2}0)$, and so

$$S(hk\ell) = \frac{1}{4}(1+e^{i\pi(k+\ell)}+e^{i\pi(h+\ell)}+e^{i\pi(h+k)})$$

$$= 1 \text{ if } h,k,\ell \text{ all odd or all even}$$

$$= 0 \text{ otherwise}$$

Diamond

The diamond lattice has eight atoms per unit cell and is simply two fcc lattices displaced by $(\frac{1}{4}\frac{1}{4}\frac{1}{4})$. The structure factor is therefore

$$S(hk\ell) = (\frac{1}{8})(1+e^{i(\frac{\pi}{2})(h+k+\ell)})(1+e^{1\pi(k+\ell)}+e^{i\pi(h+\ell)}+e^{i\pi(h+k)})$$

and is zero unless $hk\ell$ are all odd or all even. In that case, calculate $m=(4|h+k+\ell|$, modulo 4). For $m=0,1,2,3$, $|S|^2=1,\frac{1}{2},0,\frac{1}{2}$, respectively.

Hexagonal Closest Packing (hcp)

The hcp lattice has two atoms per unit cell, (000) and $(\frac{2}{3}\frac{1}{3}\frac{1}{2})$. The structure factor is

$$S(hk\ell) = \frac{1}{2}(1+e^{i(\frac{\pi}{3})(4h+2k+3\ell)})$$

calculate $m=(|4h+2k+3\ell|$, modulo 6).

If $m=0,1,2,3,4,5$; then $|S|^2=1,\frac{1}{4},\frac{3}{4},0,\frac{3}{4},\frac{1}{4}$, repectively.

Graphite

The graphite lattice has four atoms per unit cell, (000), ($\frac{2}{3}$ $\frac{1}{3}$ $\frac{1}{2}$), (00$\frac{1}{2}$), ($\frac{1}{3}$ $\frac{2}{3}$ 0). Calculate m=($|4h+2k+3\ell|$, modulo 6). If m=0,1,2,3,4,5; then $|S|^2$= 1,1/16,3/16,0, 3/16,1/16.

APPENDIX B

Numerical Evaluation

Any formula in this paper can be repetitively calculated with a short and simple digital computer program. The greatest opportunity for numerical error appears to be the integration over the finite angular spread of physical crystals. We here outline two reasonably accurate and efficient procedures that provided numbers for this paper.

Significant contributions to interference phenomena occur only for reciprocal lattice vectors $\vec{g}$ very nearly orthogonal to the photon momentum $\vec{k}$

$$\theta = \vec{g} \cdot \vec{k}/gk \approx 0 \quad .$$

One must average rapidly varying functions of this angle θ over some distribution of crystal rotations about $\vec{k} \times \vec{g}$. For one approach, we have chosen to approximate the crystal mosaic spread (p. 13) by a short series of step functions.

$$W(\theta-\bar{\theta}) \approx \sum_{n=1}^{N} a_n \Delta(\theta_n^{\,2} - (\theta-\bar{\theta})^2)$$

where here we intend $\Delta(x)$ to be the unit step function

$$\Delta(x) = 0 \qquad x < 0$$
$$\quad\;\; = 1 \qquad x > 0 \quad .$$

We then integrate the functions h_1, h_2 and d_2 (Eqs. 19, 24), defining

$$H_i(2/kg\theta) = (kg/2) \int_0^{\theta} h_i(2/kg\theta')d\theta'$$

or

$$H_i(x) = \int_x^\infty \left[\frac{n_i(x')}{(x')^2}\right]dx'$$

and similarly

$$D_2(x) = \int_x^\infty \left[\frac{d_2(x')}{(x')^2}\right]dx' \quad .$$

We find that, for $0 \le x \le 1$,

$$H_1(x) = \frac{1}{9}(14+7x+3x^2)\sqrt{1-x} - (2x+X^2 - \frac{1}{3}X^3)R$$

$$H_2(x) = \frac{1}{9}(2+X - \frac{3}{2}X^2)\sqrt{1-x} - \frac{1}{6}X^3 R$$

$$D_2(X) = \frac{1}{18\pi}[14X+3X^3(R^2+S^2-(\pi/2)^2)+(8-4x-6x^2)S\sqrt{1+x}$$
$$-(8+4x-6x^2)R\sqrt{1-x}]$$

and for $x \ge 1$.

$$H_1(x) = H_2(x) = 0$$

$$D_2(x) = \frac{1}{18\pi}[14x+3x^3(S^2-T^2)+(8-4x-6x^2)S\sqrt{x+1}$$
$$-(8+4x-6x^2)T\sqrt{x-1}]$$

where $R = 1/2\ \ln\left[(1+\sqrt{1-x})/(1-\sqrt{1-x})\right]$

$S = 1/2\ \ln\left[(\sqrt{1+x}+1)/(\sqrt{1+x}-1)\right]$

$T = \cot^{-1}(\sqrt{x-1})$.

We next observe that if $\vec{g}$ is a reciprocal lattice vector, then $-\vec{g}$ is also a reciprocal lattice vector. We can consider each such pair $(\vec{g},-\vec{g})$ only once and designate the pair by $\vec{g}$. We must then extend the functions h_1, h_2 and d_2 to negative values of $x = 2/k\cdot\vec{g}$, by requiring each of the three functions to be even functions of x and their integrals to be odd:

$$h_i(x) = h_i(-x) \qquad H_i(x) = -H_i(-x)$$
$$d_2(x) = d_2(-x) \qquad D_2(x) = -D_2(-x) \quad .$$

For numerical calculations, we first tabulate $\vec{g}$ and

$$G(\vec{g}) = (2/g)F(\vec{g}) \qquad \text{(F defined in Eq. 20b).}$$

for a few of the largest $G(\vec{g})$ with $\vec{g}$ nearly orthogonal to k. We then calculate for each of these lattice points

$$\tilde{h}_i(\vec{g}) = \sum_{j=1}^{N} a_j[H_i(2/(\vec{k}\cdot\vec{g}+kg\theta_j))-H_i(2/(\vec{k}\cdot\vec{g}-kg\theta_j))]$$

$$\tilde{d}_2(\vec{g}) = \sum_{j=1}^{N} a_j[D_2(2/(\vec{k}\cdot\vec{g}+kg\theta_j))-D_2(2/(\vec{k}\cdot\vec{g}-kg\theta_j))] \quad .$$

When the photon momentum $\vec{k}$ forms a small angle $\bar{\theta}$ with a single family of parallel crystal planes, with plane spacing d, we need consider only the reciprocal lattice vectors normal to these planes. We specify the reciprocal lattice vector pairs by magnitude

$$g_n = 2\pi n/d \qquad n > 0 \quad .$$

We may then proceed to calculate

$$\bar{\mu}_i = k \sum_n F(g_n)\ SW(\theta-\bar{\theta})h_i(2/kg_n\theta)d\theta$$

$$= \sum_n G(g_n)\tilde{h}_i(g_n)$$

and similarly

$$\overline{\Delta_2} = \sum_n G(g_n)\tilde{d}_2(g_n) \quad .$$

When $\vec{k}$ is nearly parallel to two families of parallel planes, the calculation proceeds as before except that now the calculations for $\bar{\mu}_2$ and $\bar{d}_2$ expand to

$$\bar{\mu}_2 \cos 2\phi = \sum_g G(\vec{g})\tilde{h}_2(\vec{g})\left[\frac{g_x^2 - g_y^2}{g^2}\right]$$

$$\bar{\mu}_2 \sin 2\phi = \sum_g G(\vec{g})\tilde{h}_2(\vec{g})\left[\frac{2g_x - g_y}{g^2}\right]$$

and similarly

$$\bar{\Delta}_2 \cos 2\phi' = \sum_g G(\vec{g})\tilde{d}_2(\vec{g})\left[\frac{g_x^2 - g_y^2}{g^2}\right]$$

$$\bar{\Delta}_2 \sin 2\phi' = \sum_g G(\vec{g})\tilde{d}_2(\vec{g})\left[\frac{2g_x g_y}{g^2}\right] \quad .$$

The x and y axes are arbitrarily fixed axes mutually orthogonal and transverse to $\vec{k}$. The axis of anisotropic attenuation, $\hat{t}$, defined so that attenuation is least for photons linearly polarized parallel to $\hat{t}$, has been rotated about $\vec{k}$ from the x-axis, toward the y-axis, by an angle ϕ. Similarly, the axis of anisotropic phase shift, $\hat{t}$, has been rotated away from the x-axis by ϕ'.

Note that ϕ and ϕ' are not necessarily equal. In practice, if the important $\vec{g}$ are not parallel to each other, ϕ' is not equal to ϕ, but individual contributions to $\bar{\mu}_2$ and $\bar{d}_2$ have largely cancelled.

An alternate method for calculating interference phenomena for a single family of crystal planes is to first tabulate the functions μ_1/k, μ_2/k, Δ_2/k for a large number of closely spaced values of the single variable $(k\theta)$. Numerical integration with a weight function $W(\theta-\bar{\theta})$ is then straightforward for arbitrary k and $\bar{\theta}$.

We have used both methods to calculate $\bar{\mu}$, $\bar{\mu}_2$ and $\bar{\Delta}_2$ for c-axis momentum transfer in beryllium and for graphite. Only the first method could be used for beryllium when reciprocal lattice vectors off the c-axis were included.

REFERENCES

1 H. Überall, Phys. Rev. 103, 1055 (1956); 107, 223 (1957).

2 N. Cabibbo et al, Phys. Rev. Letters 9, 270 (1962); Nuovo Cimento 27, 979 (1963).

3 G. Diambrinni Palazzi, Rev. Mod. Phys. 40, 611 (1968).

4 M.L. Ter-Mikaelian

5 R.L. Anderson et al, Phys. Rev. Letters 25, 1366 (1970).

6 R.L. Eisele et al, Nucl. Instr. Meth. 113, 489 (1973).

7 N. Cabibbo et al, Phys. Rev. Letters 9, 435 (1962).

8 W. Lee et al, NAL Expt. 87A.

9 D. Cramer, J. Waber, Acta Crystallogr. 18, 104 (1965).

10 Y. Tsai, Pair Production and Bremsstrahlung of Charged Particles, SLAC-PUB-1365 (1974).

11 R. Friedberg, Phys. Rev. A10, 21 (1974).

TABLE I

Graphite $A = 905/m^2$

hkl	g/mc	F(g)	$G(g) = \frac{2}{g} F(g)$
0 0 2	0.0072	$7.97 \cdot 10^{-7}$	$2.20 \cdot 10^{-4}$
4	0.0145	7.70	1.06
6	0.0217	4.05	0.37
8	0.0290	1.88	0.13
10	0.0362	0.84	0.05
12	0.0434	0.37	0.02

Beryllium $A = 407/m^2$

hkl	g/mc	F(g)	$G(g) = \frac{2}{g} F(g)$
0 0 2	0.0136	$5.01 \cdot 10^{-7}$	$7.40 \cdot 10^{-5}$
4	0.0271	1.48	1.10
6	0.0407	0.61	0.30
8	0.0542	0.25	0.09
0 1 0	0.0122	1.44	2.36
1	0.0140	3.59	5.13
2	0.0183	0.77	0.84
3	0.0237	1.43	1.21
4	0.0297	0.31	0.21
5	0.0360	0.62	0.34
0 2 0	0.0245	0.45	0.37
1	0.0254	1.26	0.99
2	0.0280	0.35	0.25
3	0.0318	0.81	0.51
0 3 0	0.0367	0.79	0.43
2	0.0391	0.67	0.34
4	0.0456	0.44	0.19
0 4 0	0.0489	0.09	0.04
1	0.0491	0.26	0.10

K+E 10 X 10 TO THE CENTIMETER 46 1512
18 X 25 CM. MADE IN U.S.A.
KEUFFEL & ESSER CO.

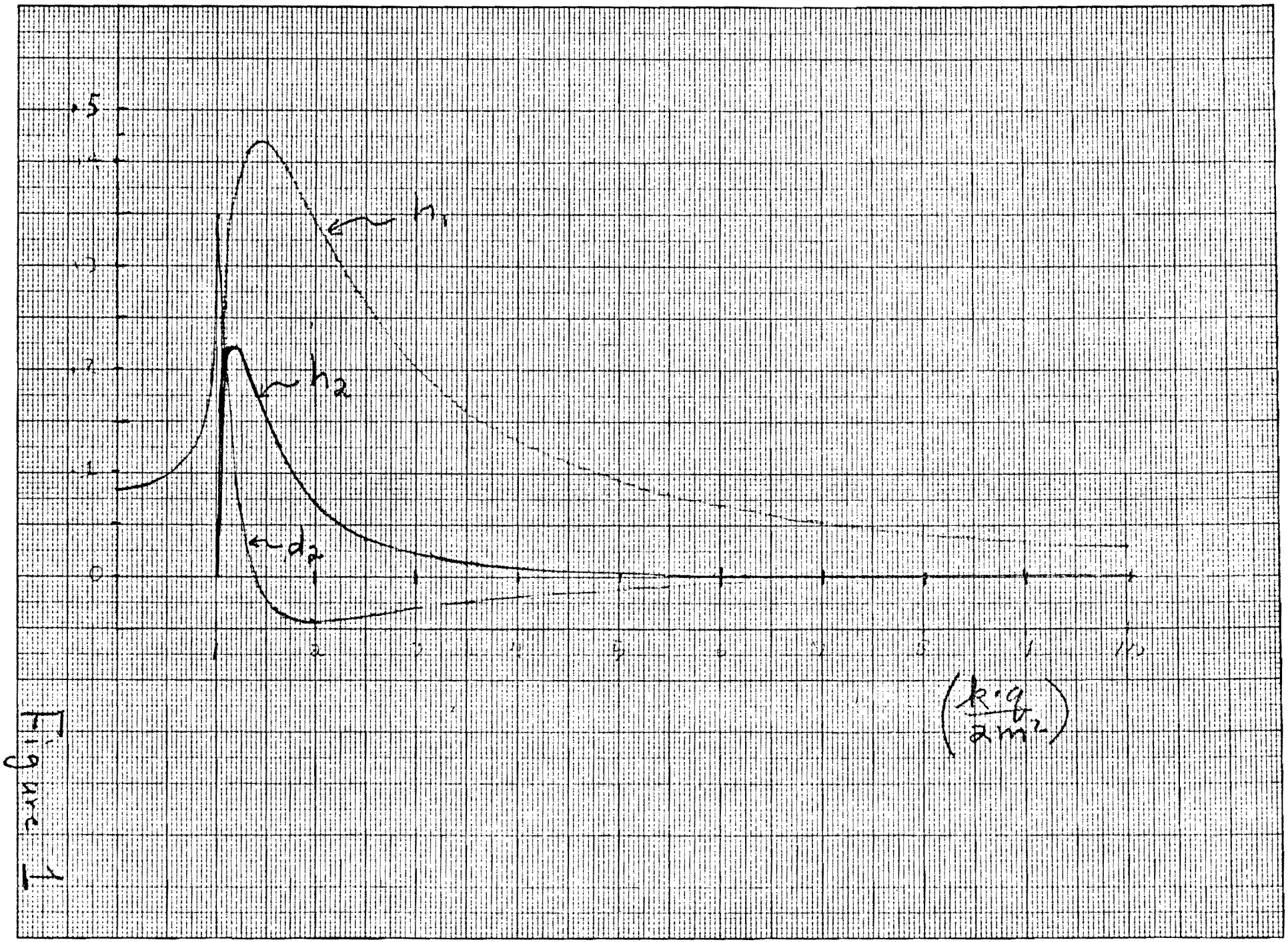
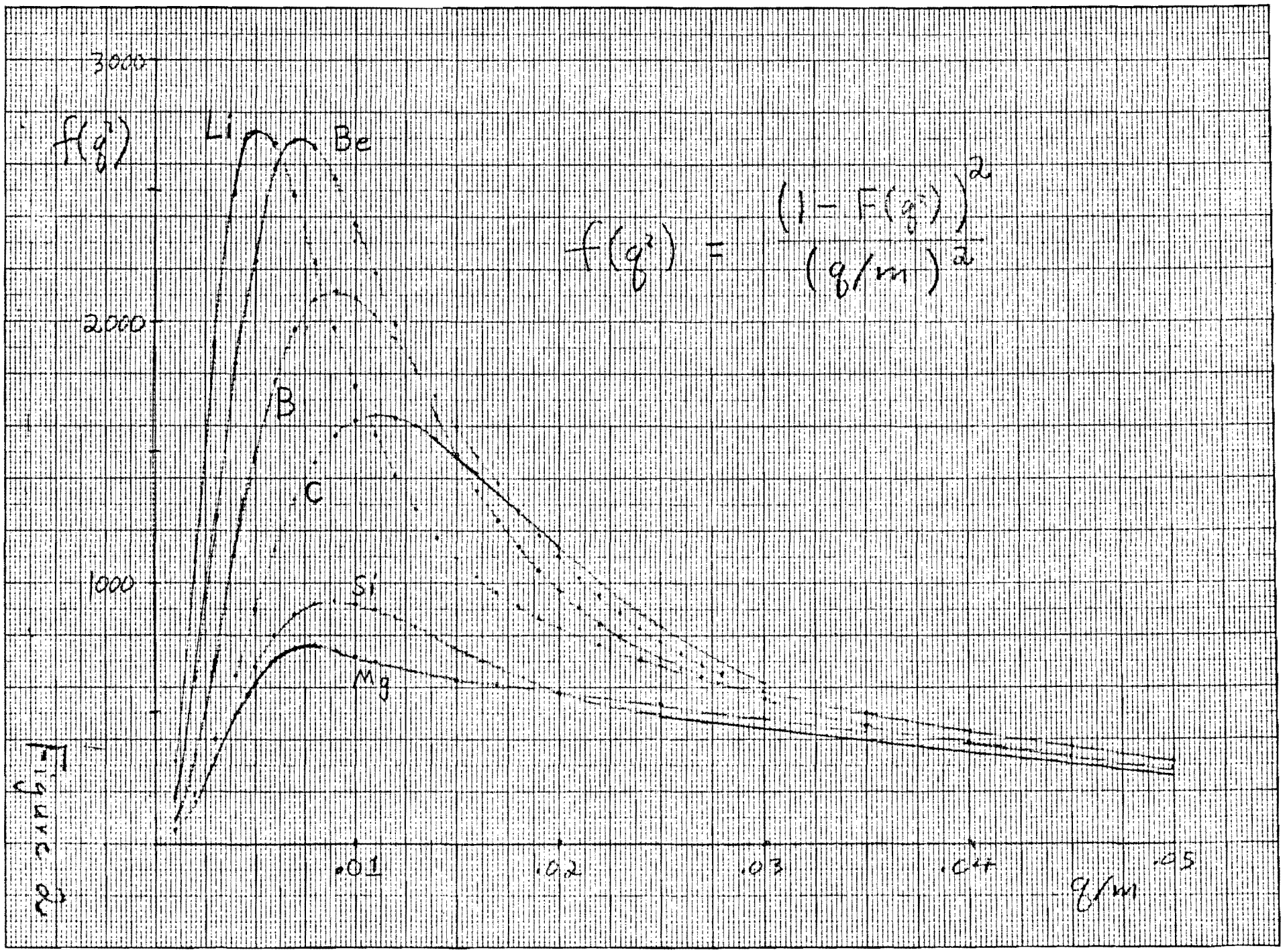


Figure 1

K+E 10 X 10 TO THE CENTIMETER 46 1512
18 X 25 CM. MADE IN U.S.A.
KEUFFEL & ESSER CO.



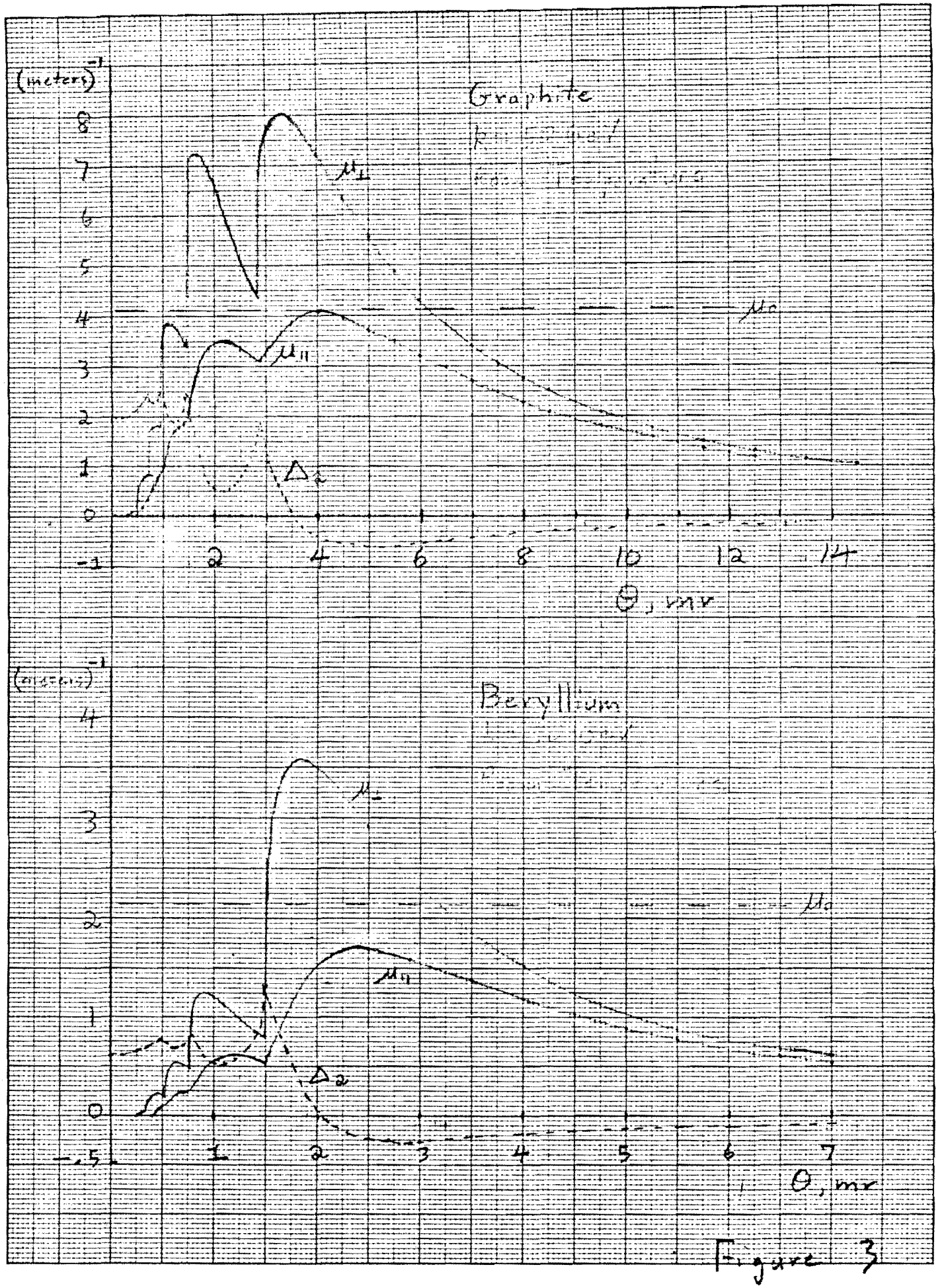
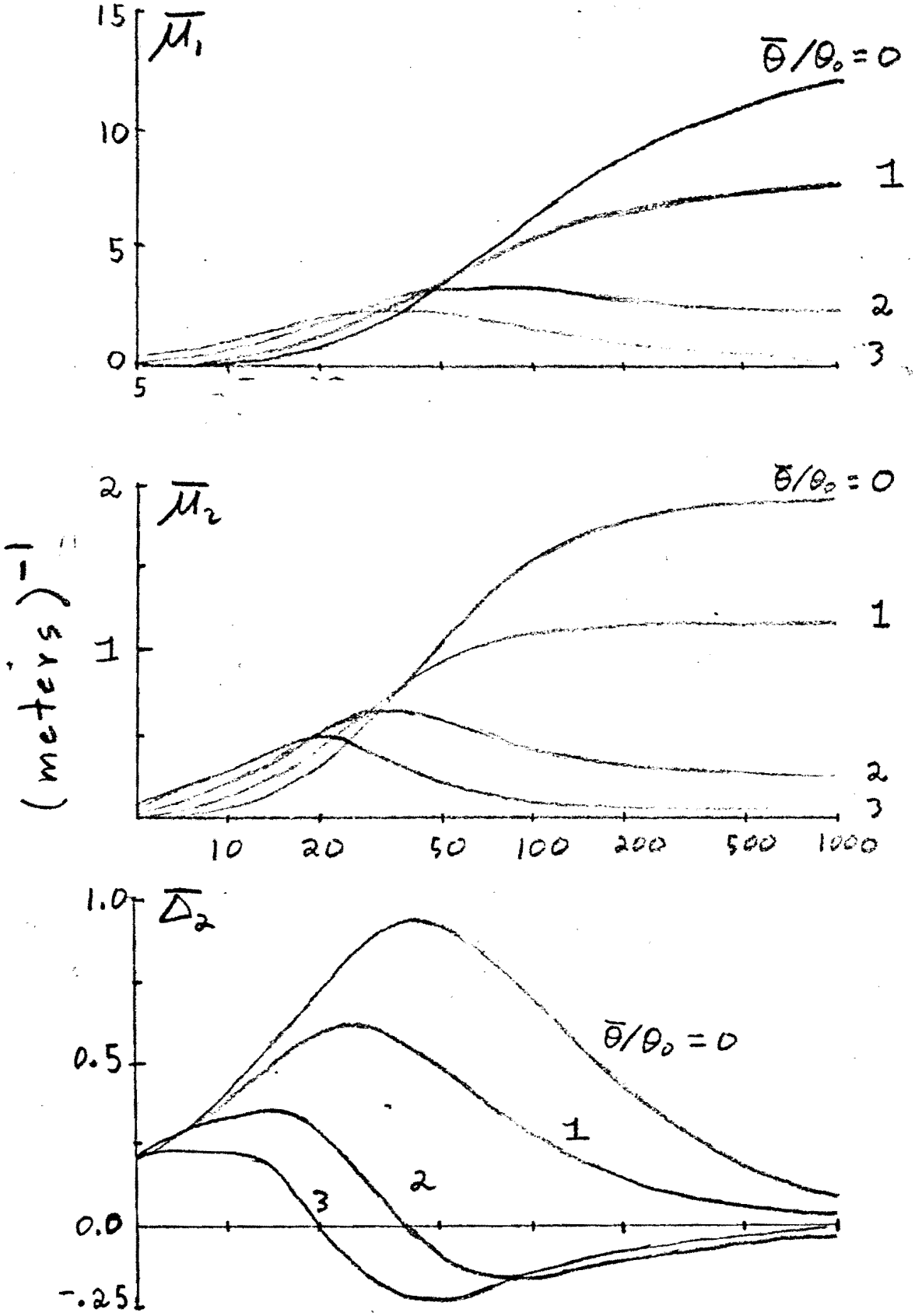


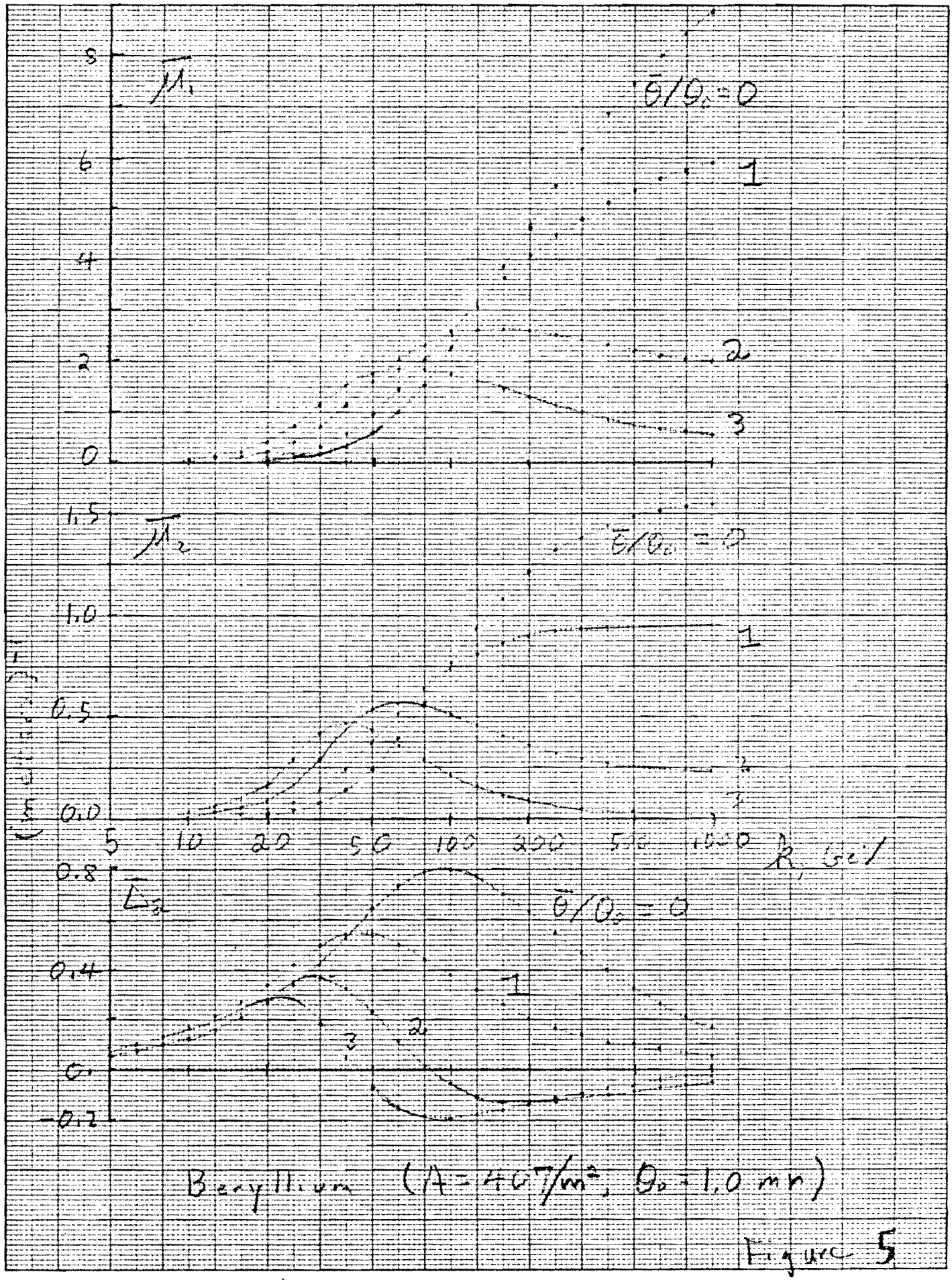
Figure 3

K+E 10 X 10 TO THE CENTIMETER 46 1512
18 X 25 CM. MADE IN U.S.A.
KEUFFEL & ESSER CO.



Graphite ($A = .905/m^2$, $\theta_0 = 3.5$ mr)

Figure 4



Beryllium ($A = 40\ T/m^2$, $\theta_0 = 1.0$ mr)

Figure 5

K+E 10 X 10 TO THE CENTIMETER 46 1512
18 X 25 CM. MADE IN U.S.A.
KEUFFEL & ESSER CO.

APPENDIX II

PHYSICAL REVIEW A VOLUME 10, NUMBER 1 JULY 1974

Creation of circular polarization by a twisted birefringent polarizing medium*

R. Friedberg
Department of Physics, Columbia University and Barnard College, New York, New York 10027
(Received 4 February 1974)

We consider the passage of initially unpolarized light through a medium possessing both birefringence and differential attenuation with respect to the same pair of perpendicular axes. Circular polarization is produced if the axes undergo rotation as the light path progresses. The twisting rate that maximizes N_2 (the circular Stokes parameter) is found to be an elliptic function with modulus depending on the path length. Various limits are discussed, and the maximized N_2 is tabulated and graphed against path length for several ratios of birefringence to differential attenuation. These results are pertinent to the polarization of hard x rays by a slightly tilted crystal.

I. INTRODUCTION

The usual way of imparting circular polarization to initially unpolarized light makes use of two media through which the light passes successively. The first medium possesses differential absorption with respect to two perpendicular directions of linear polarization. The second medium possesses differential refraction with respect to axes oriented at 45° to those of the first medium. The first medium (polarizer) transforms unpolarized into linearly polarized light. The second medium ($\lambda/4$ plate) transforms linearly into circularly polarized light.

In this paper we shall assume that only one medium is available, possessing differential absorption *and* birefringence with respect to the same pair of perpendicular directions. Mathematically, this means that the complex index of refraction is a 2×2 matrix whose eigenvectors are real and perpendicular, but whose eigenvalues are complex. Our interest in this situation arises from recent work[1-3] on the polarization of high-energy photons by crystals.

With such a medium we can create circular polarization by placing two samples of the medium one after the other, with the second rotated 45° relative to the first. The first acts as a linear polarizer; its birefringence is no embarrassment because the transmitted polarization is also on a principal axis for refraction. The second acts as a quarter-wave plate, but at the same time its absorptive properties reduce the final intensity.

As a result, the intensity of circularly polarized light (to be precise, the Stokes parameter N_2) obtainable in this way is not as great as the intensity of linear polarized light (N_3) obtained from a single untwisted sample of the same *total* length. We shall call the ratio of the former quantity to the latter the "circular conversion efficiency" or just "efficiency."

We may now vary our strategy by using any number of samples of whatever length and orientation we choose. To be utterly general, we may let the orientation of the principal axes change continuously as a function of distance measured along the path of the light. Our purpose is to choose this function so as to maximize the circular Stokes parameter N_2. For a given total path length, this is the same as maximizing the circular conversion efficiency.

In Secs. II–IV we shall set up a formalism to deal with this problem. It has two noteworthy features. First, the density matrix of the light is regarded as a four-dimensional vector $|m\rangle$. Second, a dual vector $\langle p|$ is introduced, which changes in such a way that its inner product with $|m\rangle$ is constant along the path of the light. The development of $|m\rangle$ along the path is then conveniently described in terms of various bilinear forms on $\langle p|$ and $|m\rangle$.

In Sec. V the variational problem is solved. The ordinary Euler-Lagrange method seems inapplicable because the quantity N_2 is not given as an integral and therefore the variational derivatives are not easily found. However, the formalism of the preceding sections makes it possible to show that the final N_2 is stationary with respect to changes in the twisting function, if and only if a certain bilinear form on $\langle p|$ and $|m\rangle$ vanishes everywhere along the path.

In Secs. VI and VII we use the vanishing of this form to derive a differential equation which enables us to express the optimal twisting function in terms of an elliptic function whose modulus is related indirectly to the length of the crystal. [See Eq. (73).] We do not obtain a closed expression for the density matrix at any point. However, in Secs. VIII and IX we study the maximized quantity N_2 by an indirect method and show that it depends on a complete elliptic integral of the third kind. [See Eq. (101).]

The remaining sections are devoted to subsidiary results. In Sec. X we obtain inequalities which show that the solution defined by (73) is really optimal, although an infinite set of other solutions exists which renders N_2 stationary. The limit of short path is discussed in Sec. XI, and that of long path in Sec. XII. Section XIII deals with the choice of "best" path length for a given over-all attenuation rate. In Sec. XIV we develop some numerical tables, and in Sec. XV we summarize our findings.

II. DEFINITIONS AND NOTATION

Let x be the distance from the entrance face to an arbitrary point on the light path.

Let L be the distance from the entrance face to the exit face.

Let the column vectors $\binom{1}{0}$, $\binom{0}{1}$ represent light of unit amplitude polarized along either of the two principal axes. Any combination of polarized and unpolarized light is represented, then, by a density matrix

$$M = M_0\sigma_0 + M_1\sigma_1 + M_2\sigma_2 + M_3\sigma_3, \tag{1}$$

where

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$
$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2}$$

We shall have occasion to regard M_0, M_1, M_2, M_3 also as components of a column four-vector,

$$|m\rangle = \begin{pmatrix} M_0 \\ M_1 \\ M_2 \\ M_3 \end{pmatrix}, \tag{3}$$

and we write the relation between M and $|m\rangle$ contained in (1) and (3) as

$$M \Rightarrow |m\rangle. \tag{4}$$

It will be useful also to define 4×4 matrices Σ_i^L, Σ_i^R whose effect on $|m\rangle$ corresponds to that of multiplying M either on the left or on the right by any of the σ_i. Thus if

$$A \Rightarrow |a\rangle, \tag{5}$$

then

$$\sigma_i A \Rightarrow \Sigma_i^L |a\rangle \quad (i = 1, 2, 3),$$
$$A\sigma_i \Rightarrow \Sigma_i^R |a\rangle \quad (i = 1, 2, 3). \tag{6}$$

From (6) and the σ commutation laws

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k, \tag{7}$$

we obtain the Σ commutation laws

$$[\Sigma_i^L, \Sigma_j^L] = 2i\epsilon_{ijk}\Sigma_k^L,$$
$$[\Sigma_i^R, \Sigma_j^R] = -2i\epsilon_{ijk}\Sigma_k^R, \tag{8}$$
$$[\Sigma_i^L, \Sigma_j^R] = 0.$$

We write the product of the complex index of refraction and the wave number as

$$n + n'\sigma_3 = \begin{pmatrix} n+n' & 0 \\ 0 & n-n' \end{pmatrix} \tag{9}$$

and the angle by which the axes at x (and hence our coordinate system) are rotated from their direction at the entrance face as $\phi(x)$. The rate of twisting is

$$\psi(x) = \frac{d}{dx}\phi(x). \tag{10}$$

We may note here that in the application to high-energy photons, the polarizing power of the crystal is achieved[3] by tilting a principal axis at a very small angle θ from the direction of propagation (the x direction in our treatment). The plane of this tilt is fixed in relation to the crystal axes, and it in turn determines the orientation of the principal (transverse) axes for refraction. Thus, our "twisting angle" ϕ is just the azimuth of this whole system about the direction of propagation.

III. EQUATION OF MOTION

As a result of (9) and (10), the propagation of light in a pure polarization state can be described by $\binom{a(x)}{b(x)}$, where

$$\frac{da}{dx} = i(n+n')a - b\psi,$$
$$\frac{db}{dx} = i(n-n')b + a\psi \tag{11}$$

or

$$\frac{d}{dx}\begin{pmatrix} a \\ b \end{pmatrix} = iH\begin{pmatrix} a \\ b \end{pmatrix}, \tag{12}$$

where

$$H = n\sigma_0 + n'\sigma_3 - \psi\sigma_2. \tag{13}$$

Therefore, the density matrix M develops according to

$$\frac{dM}{dx} = iHM - iMH^\dagger$$
$$= -2(\mathrm{Im}\,n)M + in'\sigma_3 M - in'^* M\sigma_3 - i\psi[\sigma_2, M]. \tag{14}$$

Equation (14) is a direct consequence of (11) when M represents pure polarized light, $M = \begin{pmatrix} aa^* & ab^* \\ ba^* & bb^* \end{pmatrix}$. If the light is not completely polarized, M cannot be written in this form. But we can always write

$M = M_1 + M_2$, where M_1 and M_2 represent two pure polarized components of random relative phase. Then the dependence of M on x is found by letting M_1 and M_2 each satisfy (14) and equating M to their sum at each x. Since (14) is linear in M, this means that M also satisfies (14) whether the light is completely polarized or not.

Using (6), we can write (14) as

$$\frac{d}{dx}|m\rangle = (-\lambda + \alpha - \psi\beta)|m\rangle, \tag{15}$$

where

$$\lambda = 2\,\mathrm{Im}n, \tag{16}$$

$$\alpha = in'\Sigma_3^L - in'^*\Sigma_3^R, \tag{17}$$

$$\beta = i\Sigma_2^L - i\Sigma_2^R. \tag{18}$$

As the initial condition for unpolarized light with unit intensity, we set

$$M(0) = \tfrac{1}{2}\sigma_0 \tag{19}$$

so that M and hence $|m\rangle$ are determined for all x by (14) or (15).

It will be convenient to introduce a vector $\langle p|$ in the dual four-dimensional space which develops with x according to

$$\frac{d}{dx}\langle p| = \langle p|(\lambda - \alpha + \psi\beta), \tag{20}$$

with the *final* condition (at $x = L$)

$$\langle p(L)|a\rangle = \mathrm{Tr}\sigma_2 A \tag{21}$$

for arbitrary $|a\rangle$ and A satisfying (5).

The circular Stokes parameter of the emergent light is given by

$$\begin{aligned} N_2 &= \mathrm{Tr}\sigma_2 M(L) \\ &= \langle p(L)|m(L)\rangle \end{aligned} \tag{22}$$

on account of (21). But from (15) and (20) we see that $\langle p(x)|m(x)\rangle$, or $\langle p|m\rangle$, is independent of x. Therefore we have

$$\begin{aligned} N_2 &= \langle p|m\rangle \\ &= \langle I\rangle. \end{aligned} \tag{23}$$

In (23) and henceforth, we use the notation

$$\langle\omega\rangle = \langle p|\omega|m\rangle \tag{24}$$

for any 4×4 matrix ω. In general, $\langle\omega\rangle$ may depend on x through the changing vectors $|m\rangle$ and $\langle p|$. In fact, from (15) and (20) we obtain

$$\frac{d}{dx}\langle\omega\rangle = \langle[\omega, \alpha - \psi\beta]\rangle \tag{25}$$

for any $\langle\omega\rangle$. If ω is the unit matrix I, the commutator vanishes, showing that $\langle I\rangle$ is constant as asserted above.

IV. COMMUTATORS

To make the use of (25) easy, we define

$$\gamma = in'\Sigma_1^L - in'^*\Sigma_1^R, \tag{26}$$

$$\alpha' = n'\Sigma_3^L + n'^*\Sigma_3^R, \tag{27}$$

$$\beta' = \Sigma_2^L + \Sigma_2^R, \tag{28}$$

$$\gamma' = n'\Sigma_1^L + n'^*\Sigma_1^R. \tag{29}$$

Using (8) with (17), (18), and (26)–(29), we obtain

$$[\alpha, \beta] = -[\beta, \alpha] = 2\gamma, \tag{30}$$

$$\begin{aligned} [\gamma, \alpha] &= 2(in'^2\Sigma_2^L - in'^{*2}\Sigma_2^R) \\ &= 2[\beta\,\mathrm{Re}(n'^2) - \beta'\,\mathrm{Im}(n'^2)], \end{aligned} \tag{31}$$

$$[\gamma, \beta] = -2\alpha, \tag{32}$$

$$[\beta', \alpha] = -2\gamma', \tag{33}$$

$$[\beta', \beta] = 0, \tag{34}$$

$$\begin{aligned} [\gamma', \alpha] &= 2(n'^2\Sigma_2^L + n'^{*2}\Sigma_2^R) \\ &= 2[\beta\,\mathrm{Im}(n'^2) + \beta'\,\mathrm{Re}(n'^2)], \end{aligned} \tag{35}$$

$$[\gamma', \beta] = -2\alpha', \tag{36}$$

$$[\alpha', \alpha] = 0, \tag{37}$$

$$[\alpha', \beta] = 2\gamma'. \tag{38}$$

From these equations we can determine $[\omega, \alpha - \psi\beta]$, where ω is any combination of $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$. We note that although there are 16 independent 4×4 matrices, these six span a closed Lie subalgebra, which is the direct sum of two three-dimensional Lie algebras as shown by (8).

V. STATIONARY CONDITION

We wish to choose the function $\psi(x)$ so that the functional N_2 will be stationary against small changes in ψ. Suppose that ψ were to be altered only in the interval $x_0 < x < x_0 + \epsilon$, where ϵ is small. Using (15), we have

$$|m(x_0+\epsilon)\rangle = |m(x_0)\rangle + \epsilon(-\lambda + \alpha - \psi\beta)|m(x_0)\rangle + O(\epsilon^2), \tag{39}$$

and since $\langle p|m\rangle$ is x independent, we may substitute $x = x_0 + \epsilon$ in the right side of (23), obtaining

$$\begin{aligned} N_2 &= \langle p(x_0+\epsilon)|m(x_0+\epsilon)\rangle \\ &= \langle p(x_0+\epsilon)|1 + \epsilon(-\lambda + \alpha - \psi\beta)|m(x_0)\rangle + O(\epsilon^2). \end{aligned} \tag{40}$$

Now, if ψ is replaced by $\psi + \delta\psi$, where $\delta\psi(x)$ vanishes for $x \leq x_0$ and for $x \geq x_0 + \epsilon$, the vectors $|m(x_0)\rangle$ and $\langle p(x_0+\epsilon)|$ will be unaltered since the boundary conditions (19) and (21) fix $|m(0)\rangle$ and

$\langle p(L)|$. Therefore N_2 will be replaced by $N_2 - \delta N_2$, where

$$\delta N_2 = \langle p(x_0+\epsilon)|\beta|m(x_0)\rangle \int_{x_0}^{x_0+\epsilon} \delta\psi(x)\,dx + O(\epsilon^2)$$
$$= \langle p(x_0)|\beta|m(x_0)\rangle \int_{x_0}^{x_0+\epsilon} \delta\psi(x)\,dx + O(\epsilon^2). \tag{41}$$

Since ϵ can be arbitrarily small and x_0 can be anything between 0 and L, we find that N_2 is stationary with respect to ψ if and only if

$$\langle\beta\rangle = 0 \tag{42}$$

for all x, in the notation of (24).

VI. DIFFERENTIAL EQUATION

From (42) and (25) we shall derive a differential equation for $\psi(x)$. We begin by observing that $\alpha - \psi\beta$ commutes with itself, and therefore

$$\frac{d}{dx}\langle\alpha\rangle - \psi\frac{d}{dx}\langle\beta\rangle = 0. \tag{43}$$

The second term on the left vanishes on account of (42). Therefore the first term is also zero and $\langle\alpha\rangle$ is a quantity independent of x.

Differentiating (42) with respect to x, and using (25) with (30), we have

$$\langle\gamma\rangle = 0 \tag{44}$$

for all x. Differentiating again and using (25) with (31) and (32), we find

$$\langle\beta\rangle\mathrm{Re}(n'^2) - \langle\beta'\rangle\mathrm{Im}(n'^2) + \psi\langle\alpha\rangle = 0 \tag{45}$$

or, in view of (42),

$$\langle\beta'\rangle = C\psi, \tag{46}$$

where C is given by

$$C = +\langle\alpha\rangle/\mathrm{Im}(n'^2) \tag{47}$$

and is x independent because of (43).

Differentiating a third time, with the help of (33) and (34), we have

$$2\langle\gamma'\rangle = -C\frac{d\psi}{dx}. \tag{48}$$

A fourth differentiation, with (35) and (36), yields

$$C\frac{d^2\psi}{dx^2} = 4[-\langle\beta\rangle\mathrm{Im}(n'^2) - \langle\beta'\rangle\mathrm{Re}(n'^2) - \langle\alpha'\rangle\psi], \tag{49}$$

into which we substitute from (42) and (46) to obtain

$$C\frac{d^2\psi}{dx^2} = 4\eta\psi, \tag{50}$$

where

$$\eta = -\langle\alpha'\rangle - C\,\mathrm{Re}(n'^2). \tag{51}$$

We now differentiate (51), using (25) with (37) and (38). The result is

$$\frac{d\eta}{dx} = 2\langle\gamma'\rangle\psi = -C\psi\frac{d\psi}{dx} \tag{52}$$

on account of (48). This equation can be integrated at once, yielding

$$\eta = -C(\tfrac{1}{2}\psi^2 + g), \tag{53}$$

where g is an undetermined constant.

Putting (53) into (50), we obtain

$$\frac{d^2\psi}{dx^2} + 2\psi^3 + 4g\psi = 0, \tag{54}$$

which is our equation for $\psi(x)$.

VII. ELLIPTIC FUNCTIONS

Multiplying (54) by $2d\psi/dx$ and integrating, we find

$$\left(\frac{d\psi}{dx}\right)^2 = -\psi^4 - 4g\psi^2 + h, \tag{55}$$

where h is another undetermined constant. Real solutions to this equation are of the form

$$\psi = \pm k\rho\,\mathrm{cn}\rho(x+l) \tag{56}$$

if $h > 0$, or

$$\psi = \pm\rho\,\mathrm{dn}\rho(x+l) \tag{57}$$

if $h < 0$. Here $\mathrm{cn}u$ and $\mathrm{dn}u$ are elliptic functions of modulus k, defined by

$$\begin{aligned} &\mathrm{cn}u = (1-\mathrm{sn}^2u)^{1/2},\\ &\mathrm{dn}u = (1-k^2\,\mathrm{sn}^2u)^{1/2},\\ &\frac{d}{du}\mathrm{sn}u = \mathrm{cn}u\,\mathrm{dn}u,\\ &\mathrm{cn}0 = \mathrm{dn}0 = 1. \end{aligned} \tag{58}$$

The quantities k and ρ depend on g and h, and l is a new undetermined constant.

We shall determine k, ρ, l and choose between (56) and (57) by examining the boundary conditions on ψ. From (6) and (19) we infer

$$\Sigma_i^L|m(0)\rangle = \Sigma_i^R|m(0)\rangle \tag{59}$$

for $i = 1, 2, 3$. Therefore, comparing (17) with (27) and (26) with (29), we have, for $x = 0$,

$$\langle\alpha'\rangle = -(\mathrm{Re}n'/\mathrm{Im}n')\langle\alpha\rangle, \tag{60}$$

$$\langle\gamma'\rangle = -(\mathrm{Re}n'/\mathrm{Im}n')\langle\gamma\rangle. \tag{61}$$

On the other hand, (6) and (21) give

$$\langle p(L)|\Sigma_i^L = -\langle p(L)|\Sigma_i^R \tag{62}$$

for $i=1$ or 3. A similar comparison yields, for $x=L$,

$$\langle\alpha'\rangle=(\mathrm{Im}n'/\mathrm{Re}n')\langle\alpha\rangle, \tag{63}$$

$$\langle\gamma'\rangle=(\mathrm{Im}n'/\mathrm{Re}n')\langle\gamma\rangle. \tag{64}$$

Substituting (44) and (48) into (61) and (64), we have

$$\frac{d\psi}{dx}=0 \tag{65}$$

for $x=0$ or L.

Substituting (47) and (60) into (51), we find

$$\begin{aligned}\eta(0)&=+(\mathrm{Re}n'/\mathrm{Im}n')\langle\alpha\rangle-C\,\mathrm{Re}(n'^2)\\&=C[(\mathrm{Re}n'/\mathrm{Im}n')\,\mathrm{Im}(n'^2)-\mathrm{Re}(n'^2)]\\&=C[2(\mathrm{Re}n')^2-(\mathrm{Re}n')^2+(\mathrm{Im}n')^2]\\&=C|n'|^2.\end{aligned} \tag{66}$$

Replacing (60) by (63), we have

$$\begin{aligned}\eta(L)&=-(\mathrm{Im}n'/\mathrm{Re}n')\langle\alpha\rangle-C\,\mathrm{Re}(n'^2)\\&=-C[(\mathrm{Im}n'/\mathrm{Re}n')\,\mathrm{Im}(n'^2)+\mathrm{Re}(n'^2)]\\&=-C[2(\mathrm{Im}n')^2+(\mathrm{Re}n')^2-(\mathrm{Im}n')^2]\\&=-C|n'|^2.\end{aligned} \tag{67}$$

Now, the derivative of $\mathrm{cn}u$ vanishes for real u only when u is a multiple of $2K$, where $4K$ is the full period of the elliptic functions on the real line:

$$K=\int_0^1\frac{dz}{(1-z^2)(1-k^2z^2)^{1/2}}. \tag{68}$$

The second derivative of $\mathrm{cn}u$ is equal to $-\mathrm{cn}u$ at all multiples of $2K$. It follows that if we use (56), then on account of (50) and (65) we must have

$$\eta(0)=\eta(L), \tag{69}$$

which contradicts (66) and (67). Therefore (56) is ruled out, and (57) is correct.

The derivative of $\mathrm{dn}u$ vanishes at all multiples of K. The second derivative of $\mathrm{dn}u$ has the value $-k^2\,\mathrm{dn}u$ at even multiples of K, and $+k^2\,\mathrm{dn}u$ at odd multiples. Applying these facts to (57), and using (65)–(67) with (50), we have

$$\rho L=(2s-1)K, \tag{70}$$

where s may be any positive integer, and

$$k\rho=2|n'|, \tag{71}$$

$$l=L. \tag{72}$$

We shall confine our attention for the time being to the *principal solution* obtained by setting $s=1$. For this solution, the combination (70) and (71) yields

$$kK=2|n'|L, \tag{73}$$

which together with (68) determines k.

From (58) we obtain

$$\int\mathrm{dn}u\,du=\int\frac{d\,\mathrm{sn}u}{\mathrm{cn}u}=\sin^{-1}(\mathrm{sn}u)+\mathrm{const.} \tag{74}$$

Thus, comparing (10) with (57), we find

$$\cos\phi=\mathrm{sn}\rho(x+L), \tag{75}$$

where we use the cosine instead of the sine in order to make $\phi(0)=0$. By combining (75) with (57) and using (58) again, we find

$$\frac{d\phi}{dx}=\psi=\pm\frac{2|n'|}{k}(1-k^2\cos^2\phi)^{1/2}, \tag{76}$$

which is a useful way to express the dependence of ϕ on x.

The sign of ψ is still undetermined. We shall show at the end of Sec. VIII how to choose it so as to make N_2 positive. Two rather surprising observations may be made. The first is that in view of (57) and (73) the whole form of $\phi(x)$ is independent of the argument of n', depending only on its absolute magnitude. The second is that regardless of the value of k we have

$$\phi(L)=\pm\pi/2 \tag{77}$$

on account of (75).

VIII. DEPENDENCE ON PATH LENGTH

In this section we shall write N_2 or $N_2(L)$ or $N_2(k)$ for the circular Stokes parameter obtained from the principal solution of (54); that is, N_2 is given by (22) with (76), where k is determined by (73) and (68). In order to compute N_2 as a function of L or k, we first study its derivative with respect to L, for fixed n and n'.

If L is replaced by $L+dL$, N_2 may be affected in two ways. First, a path length dL is added at the end. Second, ψ is slightly altered for $0<x<L$ since L enters into (57) through k, ρ, and l. But the second of these changes has no effect on N_2 to first order, since ψ has been chosen so as to make N_2 stationary. Hence, referring to (22), we have

$$\begin{aligned}\frac{dN_2}{dL}&=\mathrm{Tr}\,\sigma_2\frac{dM}{dx}\bigg|_{x=L}=\langle p(L)|\frac{d}{dx}|m(x)\rangle_{x=L}\\&=\langle p|-\lambda+\alpha-\psi\beta|m\rangle=-\lambda N_2+\langle\alpha\rangle,\end{aligned} \tag{78}$$

where we have made use of (15) and (42).

We shall now relate $\langle\alpha\rangle$ to N_2 by studying the quantity

$$|M|=M_0^2-M_1^2-M_2^2-M_3^2 \tag{79}$$

as a function of x. From (1) we have

$$M^cM=\sigma_0|M|, \tag{80}$$

where

$$M^c = \sigma_0 \operatorname{Tr} M - M = M_0\sigma_0 - M_1\sigma_1 - M_2\sigma_2 - M_3\sigma_3. \tag{81}$$

Let us define

$$A^c = \sigma_0 \operatorname{Tr} A - A \tag{82}$$

for arbitrary 2×2 matrices A. Then it is easily seen, by writing A as $A_0\sigma_0 + \vec{A}\cdot\vec{\sigma}$, that

$$(AB)^c = B^c A^c \tag{83}$$

for any A, B.

Now we combine (80) with (14) to obtain

$$\begin{aligned}\sigma_0 \frac{d}{dx}|M| &= \frac{d}{dx}(M^c M)\\ &= i[M^c H M - M^c M H^\dagger + (HM)^c M - (MH^\dagger)^c M]\\ &= i[M^c(H+H^c)M - M^c M H^\dagger - (H^\dagger)^c M^c M]\\ &= i[M^c(\operatorname{Tr} H)M - M^c M H^\dagger - (H^\dagger)^c M^c M]\\ &= i\sigma_0|M|\operatorname{Tr}(H - H^\dagger)\\ &= -2\lambda\sigma_0|M|\end{aligned} \tag{84}$$

and therefore, using (19),

$$4|M| = e^{-2\lambda x}. \tag{85}$$

At $x = L$ we can apply (21) to obtain

$$\begin{aligned}2M_0 &= \langle\Sigma_2^L\rangle = \langle\Sigma_2^R\rangle = \tfrac{1}{2}\langle\beta'\rangle,\\ 2M_1 &= -i\langle\Sigma_3^L\rangle = i\langle\Sigma_3^R\rangle = -\frac{1}{2\operatorname{Re} n'}\langle\alpha\rangle,\\ 2M_2 &= \langle I\rangle = N_2,\\ 2M_3 &= i\langle\Sigma_1^L\rangle = -i\langle\Sigma_1^R\rangle = \frac{1}{2\operatorname{Re} n'}\langle\gamma\rangle.\end{aligned} \tag{86}$$

Substituting from (44), (46), and (47), we find

$$4|M| = Q\langle\alpha\rangle^2 - N_2^2, \tag{87}$$

where

$$\begin{aligned}Q &= \frac{1}{4}\left(\frac{\psi(L)^2}{\operatorname{Im}(n'^2)^2} - \frac{1}{(\operatorname{Re} n')^2}\right)\\ &= \frac{1}{\operatorname{Im}(n'^2)^2}\left[\tfrac{1}{4}\psi(L)^2 - (\operatorname{Im} n')^2\right]\\ &= \frac{1}{\operatorname{Im}(n'^2)^2}\left[\tfrac{1}{4}\rho^2 - (\operatorname{Im} n')^2\right]\\ &= \frac{1}{\operatorname{Im}(n'^2)^2}\left[\frac{|n'^2|}{k^2} - (\operatorname{Im} n')^2\right].\end{aligned} \tag{88}$$

Here we have used (57) and (71), remembering that $\operatorname{dn} 2K = 1$.

Comparing (87) with (85) for $x = L$, we have

$$Q\langle\alpha\rangle^2 = N_2^2 + e^{-2\lambda L}. \tag{89}$$

But if we write

$$N_2 = \mathfrak{F}(L)e^{-\lambda L} \tag{90}$$

and substitute in both (89) and (78), we find

$$\frac{d\mathfrak{F}}{dL} = \langle\alpha\rangle e^{\lambda L} = [Q^{-1}(1+\mathfrak{F}^2)]^{1/2} \tag{91}$$

and hence

$$\mathfrak{F} = \sinh y, \tag{92}$$

where

$$\frac{dy}{dL} = \frac{1}{Q^{1/2}}. \tag{93}$$

The sign of y is that of $Q^{1/2}$, which by (91) is that of $\langle\alpha\rangle$. Thus, $y, \mathfrak{F}, N_2$ will be positive if $\langle\alpha\rangle$ is. That means, by (46) and (47), that we should give ψ the same sign as $\langle\beta'\rangle\operatorname{Im}(n'^2)$. The sign of ψ is the same for all x, and $\langle\beta'\rangle$ must be positive at $x = L$ by (86). Therefore the sign in (57) should be taken as that of $\operatorname{Im}(n'^2)$ or of $\operatorname{Re} n' \operatorname{Im} n'$, if N_2 is to be positive. The physical sense of this may be inferred from (11) and (22) with (2).

IX. ELLIPTIC INTEGRAL

Let us now write

$$n' = |n'|e^{i\chi} \tag{94}$$

and

$$q = (1 - k^2\sin^2\chi)^{1/2}. \tag{95}$$

Then (88) becomes

$$Q = q^2/k^2|n'^2|\sin^2 2\chi, \tag{96}$$

which with (93) and (73) gives, taking ψ positive,

$$\frac{dy}{dk} = \frac{\sin 2\chi}{2}\,\frac{k}{q}\,\frac{d}{dk}(kK). \tag{97}$$

It will now be prudent to replace (68) by a contour integral:

$$K = \tfrac{1}{4}\oint \frac{dz}{(1-z^2)(1-k^2z^2)^{1/2}}, \tag{98}$$

where the contour runs counterclockwise around the cut from -1 to 1, and the square roots are both positive when z is negative imaginary. We then have

$$y = \frac{\sin 2\chi}{8}\oint \frac{dz}{(1-z^2)^{1/2}} f(k,z), \tag{99}$$

where

$$f(k,z)=\int\frac{k}{q}\frac{d}{dk}\frac{k}{(1-k^2z^2)^{1/2}}\,dk$$
$$=\int(1-k^2z^2)^{-3/2}\frac{k\,dk}{q}$$
$$=-\frac{1}{\sin^2\chi}\int(1-k^2z^2)^{-3/2}\,dq$$
$$=-\frac{1}{\sin^2\chi}\int\left(1-\frac{z^2}{\sin^2\chi}+\frac{q^2z^2}{\sin^2\chi}\right)^{-3/2}dq$$
$$=-\frac{1}{\sin^2\chi}\left(1-\frac{z^2}{\sin^2\chi}\right)^{-1}$$
$$\times q\left(1-\frac{z^2}{\sin^2\chi}+\frac{q^2z^2}{\sin^2\chi}\right)^{-1/2}$$
$$=\frac{1}{z^2-\sin^2\chi}\left(\frac{1-k^2\sin^2\chi}{1-k^2z^2}\right)^{1/2},\qquad(100)$$

and so

$$N_2=e^{-\lambda L}\sinh y$$
$$=e^{-\lambda L}\sinh\left(\frac{\sin2\chi}{8}(1-k^2\sin^2\chi)^{1/2}\right.$$
$$\left.\times\oint\frac{dz}{(z^2-\sin^2\chi)(1-z^2)^{1/2}(1-k^2z^2)^{1/2}}\right).\qquad(101)$$

To show that we have not lost a constant of integration, we remark that when $L\to0$ we have $k\to0$, so that the branch points at $\pm1/k$ recede to ∞. Then the contour can be expanded and we get $y\to0$, hence $N_2\to0$ as desired.

X. INEQUALITIES

By comparison with (101), the linear polarization intensity obtained from a constant-θ sample of length L is

$$N_3=e^{-\lambda L}\sinh y_0=e^{-\lambda L}\sinh(2|n'|L\sin\chi)\qquad(102)$$

and the circular conversion efficiency is therefore

$$E=\sinh y/\sinh y_0.\qquad(103)$$

From (102) and (97) and (73), we find

$$\frac{dy}{dy_0}=\frac{k\cos\chi}{(1-k^2\sin^2\chi)^{1/2}}<1\qquad(104)$$

so that $y<y_0$ and $E<1$, as expected.

We shall now show that the principal solution gives the highest efficiency for a given L. First, we note that if we vary the integer s in (70), the equations of Sec. VIII are completely unaffected. However, instead of (73) we have

$$(2s-1)kK=2|n'|L\qquad(105)$$

and this will affect the equations of Sec. IX.

Let us define y_s as the solution of (93) with (88) and (105). Then instead of (97) we have

$$y_s=\tfrac{1}{2}\sin2\chi\,(2s-1)\,Y(k_s),\qquad(106)$$

where Y is the function satisfying

$$\frac{dY(k)}{dk}=\frac{k}{q}\frac{d}{dk}Z(k)\qquad(107)$$

and $Y(0)=0$, and k_s is the value of k satisfying

$$(2s-1)Z(k_s)=2|n'|L,\qquad(108)$$

with the function Z given by

$$Z(k)=kK.$$

It is always understood that K and q depend on k through (68) and (95).

Combining (106) with (108), we have

$$y_s=[Y(k_s)/Z(k_s)]\,|n'|L\sin2\chi.\qquad(109)$$

Since Z is an increasing function of k, we have from (108) the inequality

$$k_1>k_2>k_3>\cdots.\qquad(110)$$

From (109), therefore, we obtain

$$y_1>y_2>y_3>\cdots\qquad(111)$$

provided that Y/Z is also an increasing function of k.

To prove the latter assertion, we write (107) as

$$\frac{dY}{dZ}=\frac{k}{q},\qquad(112)$$

and considering Z as the independent variable, we apply the "law of the mean" to the interval from 0 to Z. The result is

$$Y/Z=k'/q',\qquad(113)$$

where k' lies between 0 and k. Since k/q is obviously an increasing function of k, we may compare (113) with (112) to obtain

$$\frac{dY}{dZ}>\frac{Y}{Z},\qquad(114)$$

which leads easily to the desired inequality,

$$\frac{d}{dk}\left(\frac{Y}{Z}\right)>0\qquad(115)$$

for $0<k<1$. This shows that Y/Z is an increasing function and completes the proof of (111).

Applying (111) to (90) and (92), we see that the efficiency at a given L is highest for the principal solution ($s=1$) and decreases with increasing s.

XI. SHORT-PATH LIMIT

When $|n'|L\ll1$, we have $k\ll1$ and $K\approx\pi/2$, and (57) reduces to (taking the positive sign)

$$\psi=\rho=\pi/2L\qquad(116)$$

for the principal solution, or

$$\phi(x)=\tfrac{1}{2}\pi x/L. \tag{117}$$

The leading term of the integral in (101) is then

$$\oint \frac{dz}{(z^2-\sin^2\chi)(1-z^2)^{1/2}}\,\frac{k^2z^2}{2}$$
$$=\frac{k^2}{2}\left(\oint\frac{dz}{(1-z^2)^{1/2}}+\oint\frac{\sin^2\chi\,dz}{(z^2-\sin^2\chi)(1-z^2)^{1/2}}\right)$$
$$=\pi k^2 \tag{118}$$

and therefore

$$N_2\simeq e^{-\lambda L}y\simeq\tfrac{1}{8}\pi k^2\sin2\chi\,e^{-\lambda L}$$
$$=\frac{\pi|n'|^2L^2\sin2\chi}{2(\pi/2)^2}\,e^{-\lambda L}$$
$$=(2/\pi)|n'|^2L^2\sin2\chi\,e^{-\lambda L}$$
$$=(4/\pi)L^2\,\mathrm{Re}n'\,\mathrm{Im}n'\,e^{-\lambda L}\,. \tag{119}$$

As a check, we shall derive (117) and (119) directly from (14) in the limit

$$g\equiv|n'|L\ll1. \tag{120}$$

Setting

$$r=x/L \tag{121}$$

and

$$\xi=\frac{d\phi}{dr} \tag{122}$$

we write (14) as

$$\frac{dM}{dr}=-i\xi[\sigma_2,M]-\lambda LM+ig(e^{i\chi}\sigma_3M-e^{-i\chi}M\sigma_3), \tag{123}$$

and using (19) and (22) we obtain (90), where $\mathfrak{F}$ can be expanded in powers of g,

$$\mathfrak{F}=\mathfrak{F}_0+g\mathfrak{F}_1+g^2\mathfrak{F}_2+\cdots. \tag{124}$$

A straightforward iterative solution of (123), with the notation

$$U(r_2,r_1)=\exp\left(-i\sigma_2\int_{r_1}^{r_2}\xi\,dr\right)=e^{-i\sigma_2[\phi(r_2)-\phi(r_1)]}, \tag{125}$$

yields

$$\mathfrak{F}_0=\tfrac{1}{2}\,\mathrm{Tr}\{\sigma_2U(1,0)\sigma_0U^{-1}(1,0)\}=\tfrac{1}{2}\,\mathrm{Tr}\sigma_2=0, \tag{126}$$

$$\mathfrak{F}_1=\tfrac{1}{2}\,\mathrm{Tr}\left\{\sigma_2\int_0^1dr_1\,U(1,r_1)ie^{i\chi}\sigma_3\times U(r_1,0)\sigma_0U^{-1}(1,0)\right\}+\text{c.c.}$$
$$=\mathrm{Re}\,ie^{i\chi}\int_0^1dr_1\,\mathrm{Tr}\{\sigma_2U(1,r_1)\sigma_3U^{-1}(1,r_1)\}$$
$$=\mathrm{Re}\,ie^{i\chi}\,\mathrm{Tr}(\sigma_2\sigma_3)=0, \tag{127}$$

$$\mathfrak{F}_2=\int_0^1dr_1\int_{r_1}^1dr_2[F(r_1,r_2)+F'(r_1,r_2)]+\text{c.c.}, \tag{128}$$

where

$$F(r_1,r_2)=\tfrac{1}{2}\,\mathrm{Tr}\{\sigma_2U(1,r_2)ie^{i\chi}\sigma_3U(r_2,r_1)ie^{i\chi}\sigma_3\times U(r_1,0)\sigma_0U^{-1}(1,0)\}$$
$$=-\tfrac{1}{2}e^{2i\chi}\,\mathrm{Tr}\{\sigma_2\sigma_3U(r_2,r_1)\sigma_3U^{-1}(r_2,r_1)\}$$
$$=-\tfrac{1}{2}e^{2i\chi}\,\mathrm{Tr}\{\sigma_2U^{-2}(r_2,r_1)\}$$
$$=-ie^{2i\chi}\sin2[\phi(r_2)-\phi(r_1)] \tag{129}$$

and

$$F'(r_1,r_2)=\tfrac{1}{2}\,\mathrm{Tr}\{\sigma_2U(1,r_2)ie^{i\chi}\sigma_3U(r_2,0)\sigma_0U^{-1}(r_1,0)\times(-ie^{-i\chi})\sigma_3U^{-1}(1,r_1)\}$$
$$=\tfrac{1}{2}\,\mathrm{Tr}\{\sigma_2\sigma_3U(r_2,r_1)\sigma_3U^{-1}(r_2,r_1)\}$$
$$=i\sin2[\phi(r_2)-\phi(r_1)]. \tag{130}$$

Substituting (129) and (130) into (128), we have

$$\mathfrak{F}_2=2\sin2\chi\int_0^1dr_1\int_{r_1}^1dr_2\sin2[\phi(r_2)-\phi(r_1)]. \tag{131}$$

Our task now is to choose the function $\phi(r)$ so as to maximize the integral in (131). For reasons shortly to appear, we define a function f on the interval $(0,2\pi)$ by

$$f(\mu)=\tfrac{1}{2}\int_0^1\delta(|\sin[\tfrac{1}{2}\mu-2\phi(r)]|)\,dr=\tfrac{1}{4}\sum\left|\frac{d\phi}{dr}\right|^{-1}, \tag{132}$$

where the sum goes over those r satisfying

$$\phi(r)\equiv\tfrac{1}{4}\mu,\quad \mathrm{mod}\,\pi/2. \tag{133}$$

Thus $f(\mu)$ is determined by $\phi(r)$, and satisfies

$$\int_0^{2\pi}f(\mu)\,d\mu=1 \tag{134}$$

as seen from the first line of (132). Moreover, we have

$$\mathfrak{F}_2\leq2\sin2\chi\int_0^1dr_1\int_{r_1}^1dr_2|\sin2[\phi(r_2)-\phi(r_1)]|$$
$$=\sin2\chi\int_0^{2\pi}f(\mu_1)\,d\mu_1\int_0^{2\pi}f(\mu_2)\,d\mu_2|\sin\tfrac{1}{2}(\mu_2-\mu_1)|. \tag{135}$$

On the other hand, given f satisfying (134), we can choose ϕ satisfying (132) so that equality is achieved in (135). This is done by defining $\phi(r)$ through the equation

$$\frac{d\phi}{dr}=1/4f(4\phi), \tag{136}$$

with $\phi(0)=0$, so that ϕ increases monotonically but never exceeds $\pi/2$, and $\sin 2[\phi(r_2)-\phi(r_1)]$ is never negative.

To maximize (131), therefore, we must choose f so as to maximize the second line of (135) with the restriction (134). Since $2|\sin\frac{1}{2}(\mu_2-\mu_1)|$ is just the length of the chord joining the points μ_1 and μ_2 on the unit circle, we may regard f as the density of matter distributed on the unit circle, and repelling itself with a force whose magnitude between two given point masses is independent of their separation. Then the total mass is 1 by (134), and the total potential energy is proportional to minus the second line of (135). This energy is minimized in the position of stable equilibrium, which intuition may identify as the uniform distribution of matter around the circle.

To confirm this intuition, we write

$$f(\mu)=\frac{1}{2\pi}+\sum_1^\infty (a_m e^{im\mu}+a_m^* e^{-im\mu}) \tag{137}$$

and note that the second line of (135) becomes

$$\sin 2\chi\left(\frac{A_0}{2\pi}+2\sum_1^\infty |a_m|^2 A_m\right), \tag{138}$$

where

$$A_m=\int_0^{2\pi}\sin\tfrac{1}{2}\mu\cos m\mu\, d\mu=\frac{1}{m+\frac{1}{2}}-\frac{1}{m-\frac{1}{2}} \tag{139}$$

which is negative for $m>0$. Therefore (138) is maximized by setting all $a_m=0$. Now, using (137) and (136), we get

$$\phi(r)=(\pi/2)r, \tag{140}$$

which in view of (121) is the same as (117).

The maximum $\mathfrak{F}_2$ is obtained from (138) and (139):

$$\mathfrak{F}_2=(A_0/2\pi)\sin 2\chi=(2/\pi)\sin 2\chi, \tag{141}$$

and when this is put into (124) and the result combined with (90) and (120), we recover (119) for the leading order in g.

XII. LONG-PATH LIMIT

When $|n'|L\gg 1$, we have $k\to 1$ and $K\sim|n'|L$ for the principal solution. In this limit the integral in (101) blows up, as the contour is pinched between two branch points. But this is expected, as y should behave like y_0 as defined by (102). Combining (101) with (98), and using (73), we have

$$y\frac{(1-k^2\sin^2\chi)^{1/2}}{\cos\chi}-ky_0=\tfrac{1}{4}\sin\chi\left((1-k^2\sin^2\chi)\oint\frac{dz}{(1-z^2)^{1/2}(1-k^2z^2)^{1/2}(z^2-\sin^2\chi)}-k^2\oint\frac{dz}{(1-z^2)^{1/2}(1-k^2z^2)^{1/2}}\right)$$
$$=\tfrac{1}{4}\sin\chi\oint\frac{dz}{z^2-\sin^2\chi}\,\frac{(1-k^2z^2)^{1/2}}{(1-z^2)^{1/2}}, \tag{142}$$

in which both sides approach a finite limit as $k\to 1$.

The quantities y and y_0 grow only logarithmically in $1-k$; therefore $O(y(1-k),y_0(1-k))$ can be neglected and the left side of (142) replaced by $y-y_0$.

On the right side, the factor $(1-k^2z^2)^{1/2}/(1-z^2)^{1/2}$ becomes +1 on the negative imaginary side of the cut, and −1 on the positive imaginary side. The limiting form of (142) is therefore

$$\begin{aligned}y-y_0&\cong\tfrac{1}{4}\sin\chi\, 2P\int_{-1}^{1}\frac{dz}{z^2-\sin^2\chi}\\&=\tfrac{1}{4}\left(P\int_{-1}^{1}\frac{dz}{z-\sin\chi}-P\int_{-1}^{1}\frac{dz}{z+\sin\chi}\right)\\&=\tfrac{1}{4}\left(\ln\frac{1-\sin\chi}{1+\sin\chi}-\ln\frac{1+\sin\chi}{1-\sin\chi}\right)\\&=\tfrac{1}{2}\ln\frac{1-\sin\chi}{1+\sin\chi}\end{aligned} \tag{143}$$

and so, from (103), the limiting efficiency is given by

$$E\to\left(\frac{1-\sin\chi}{1+\sin\chi}\right)^{1/2}=\frac{\cos\chi}{1+\sin\chi}. \tag{144}$$

This efficiency is achieved by the limiting form of (57),

$$\psi=2|n'|\operatorname{sech}2|n'|(L-x), \tag{145}$$

or of (75),

$$\cos\phi=\tanh 2|n'|(L-x). \tag{146}$$

Let us compare (144) with the result of a cruder strategy, in which we put

$$\begin{aligned}\phi(x)&=0,\quad x<L-x_0\\&=\phi_0,\quad x>L-x_0.\end{aligned} \tag{147}$$

It is easily found, for fixed x_0 and $L\to\infty$, that

$$E\to 2e^{-2x_0|n'|\sin\chi}\sin\phi_0\cos\phi_0\sin(2x_0|n'|\cos\chi), \tag{148}$$

which is maximized by setting $\phi_0=\pi/4$, $2x_0|n'|\cos\chi=\pi/2-\chi$. We then have

$$E\to\cos\chi\, e^{-(\pi/2-\chi)\tan\chi}. \tag{149}$$

Comparing the crude efficiency E_{cr}, from (149), with the ideal efficiency E_I, from (144), we have

$$E_{cr}/E_I = (1+\sin\chi)e^{-(\pi/2-\chi)\tan\chi}$$
$$\sim 1-(\pi/2-1)\chi, \quad \chi \ll 1 \quad (\mathrm{Im}n' \ll \mathrm{Re}n')$$
$$= (1+\tfrac{1}{2}\sqrt{2})\,e^{-\pi/4} \sim 0.778, \quad \chi = \pi/4 \quad (\mathrm{Im}n' = \mathrm{Re}n')$$
$$\to 2/e \sim 0.736, \quad \chi \to \pi/2 \quad (\mathrm{Im}n' \gg \mathrm{Re}n'). \tag{150}$$

XIII. OPTIMUM LENGTH

To maximize N_2 with respect to L, we differentiate (90) with the help of (92) and (93), obtaining

$$\frac{dN_2}{dL} = -\lambda e^{-\lambda L}\sinh y + \frac{1}{Q^{1/2}}\,e^{-\lambda L}\cosh y, \tag{151}$$

which vanishes when

$$Q^{1/2}\tanh y = 1/\lambda. \tag{152}$$

If $\lambda \gg |n'|$, we may simplify (152) by using the short-path limit—see (119)—which gives

$$\frac{1}{\lambda} \sim \frac{qy}{k|n'|\sin 2\chi} \sim \frac{\pi k}{8|n'|} \sim \frac{L}{2}, \tag{153}$$

and thus

$$N_2 \simeq \frac{8}{\pi}\left(\frac{|n'|}{\lambda e}\right)^2 \sin 2\chi = \frac{4}{\pi e^2}\,\frac{\mathrm{Re}n'\,\mathrm{Im}n'}{(\mathrm{Im}n)^2} \tag{154}$$

for the maximum circular intensity attainable when the over-all attenuation is much stronger than the polarizing effects.

The opposite limit occurs when one component is not attenuated at all; $\mathrm{Im}(n-n')=0$, or $\lambda = 2|n'|\sin\chi$. Combining (152) with (96), we then have

$$\tanh y = (k/q)\cos\chi, \tag{155}$$

which is satisfied when $k \to 1$, $y \to \infty$. Thus the best result is now obtained with $L \to \infty$. Since the untwisted crystal in this case will give $N_3 = \frac{1}{2}$, we have from (144)

$$N_2 \to \cos\chi/2(1+\sin\chi) \tag{156}$$

for the maximum circular intensity in this case.

TABLE I. For eight values of k we tabulate K, and the three quantities y_0, y, E for each of two values of χ ($\tan\chi = \frac{1}{2}, 3$). The tabulated values of y_0 and E are related to the two outer curves in Fig. 1.

		$\mathrm{Im}n' = \frac{1}{2}\mathrm{Re}n'$			$\mathrm{Im}n' = 3\,\mathrm{Re}n'$		
k	K	y_0	y	E	y_0	y	E
0.1	1.57	0.070	0.003	0.045	0.149	0.002	0.016
0.5	1.69	0.377	0.088	0.229	0.800	0.070	0.063
0.8	2.00	0.714	0.299	0.390	1.514	0.269	0.126
0.9	2.28	0.918	0.468	0.460	1.947	0.471	0.142
0.94	2.51	1.054	0.591	0.496	2.236	0.644	0.149
0.98	3.02	1.324	0.849	0.546	2.809	1.078	0.157
0.995	3.70	1.645	1.165	0.580	3.490	1.694	0.160
1	∞	∞	∞	0.618	∞	∞	0.162

XIV. NUMERICAL EVALUATION

We seek numerical values of the circular conversion efficiency (see Secs. I and X) given by

$$E = \sinh y/\sinh y_0, \tag{103}$$

where, from (102),

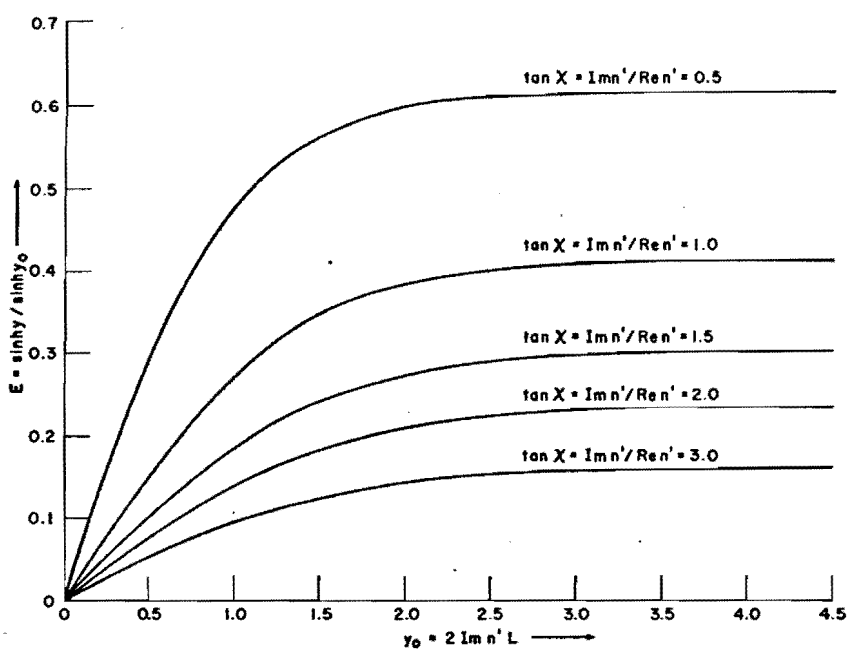


FIG. 1. Maximum possible circular conversion efficiency E (see Sec. X) is graphed against the crystal thickness, measured by the parameter y_0, for five values of $\tan\psi = \mathrm{Im}n'/\mathrm{Re}n'$. The top curve represents strong birefringence; the bottom curve represents strong polarization. The *fractional* linear polarization obtained from an *untwisted* sample is $\tanh y_0 = 0.76$ ($y_0 = 1$), $= 0.96$ ($y_0 = 2$).

$$y_0 = 2\,\mathrm{Im}\,n'L = kK\sin\chi \tag{157}$$

and y is given by (97) or (101).

It is easy to find K for a given k by the formulas

$$K = \frac{\pi}{2r_\infty}, \quad r_\infty = \lim_{m\to\infty} r_m, \quad r_{m+1} = \tfrac{1}{2}(r_m + s_m),$$
$$s_{m+1} = (r_m s_m)^{1/2}, \quad r_0 = (1-k^2)^{1/2}, \quad s_0 = 1. \tag{158}$$

This yields y_0 immediately for given k and χ.

The tabulation of y is most easily accomplished by integrating (97) by parts:

$$\frac{d}{dk}\left(\frac{y_0 k\cos\chi}{(1-k^2\sin^2\chi)^{1/2}} - y\right) = \frac{y_0\cos\chi}{(1-k^2\sin^2\chi)^{3/2}}. \tag{159}$$

The right side is easily calculated for given k and χ, and hence it is a suitable integrand for numerical integration. The integrand diverges only logarithmically at $k\to 1$, so that the integral is convergent there. (It should be understood that χ is fixed during the integration which is begun at $y = y_0 = k = 0$.)

Values of k, y_0, y, and E are tabulated for various χ in Table I. The relation between E and y_0, for fixed χ, is displayed in Fig. 1.

XV. SUMMARY OF RESULTS

We have found that the circular Stokes parameter N_2 is maximized for a given medium and path length when the twisting rate $\psi = d\phi/dx$ is given by

$$\psi(x) = \pm\frac{2|n'|}{k}\,\mathrm{dn}\frac{2|n'|}{k}(x+L), \tag{160}$$

where dn is a standard elliptic function of modulus k, and k is determined by

$$kK = 2|n'|L, \tag{73}$$

where K is the quarter-period associated with k,

$$K = \int_0^1 \frac{dz}{(1-z^2)^{1/2}(1-k^2z^2)^{1/2}}, \tag{68}$$

L is the path length, and $2n'$ is the difference in complex wave numbers between the two perpendicular principal axes.

An equivalent form of the solution is the differential equation

$$\frac{d\phi}{dx} = \pm\frac{2|n'|}{k}(1-k^2\cos^2\phi)^{1/2}. \tag{76}$$

The sign of N_2 is the same as that of $\psi\,\mathrm{Re}\,n'\,\mathrm{Im}\,n'$. The relation between signs can best be understood by considering the circular polarization produced by the standard 45° configuration, $\psi = \pm\frac{1}{4}\pi\delta(x-x_0)$.

The solution (160) and (73) gives the biggest possible N_2, although the variation of N_2 is also zero if the left side of (73) is multiplied by any odd integer.

The optimum value of N_2 is given by

$$N_2 = e^{-\lambda L}\sinh y, \tag{161}$$

where $\lambda = 2\,\mathrm{Im}\,n$ is the common attenuation rate (average of two axes) and y is a somewhat complicated function of k and $\chi = \arg n'$, given by (97) or (101).

We compare this with the linear Stokes parameter N_3 produced by an untwisted sample of the same medium with the same length,

$$N_3 = e^{-\lambda}\sinh y_0, \tag{162}$$

where

$$y_0 = 2\,\mathrm{Im}\,n'L. \tag{163}$$

The ratio N_2/N_3 does not involve λ, and is given by

$$E = \sinh y/\sinh y_0, \tag{103}$$

which is a function of k and χ, or of y_0 and χ. It is plotted in Fig. 1 against y_0 for several values of χ.

When $y_0 \ll 1$, N_2 is maximized by a uniform twisting rate, and is $O[y_0^2]$ so that E is $O[y_0]$. When $y_0 \gg 1$, the optimum strategy puts nearly all the twisting near the exit face, with a hyperbolic secant form for ψ. In this limit, E approaches the constant value $\cos\chi/(1+\sin\chi)$.

In all cases the total twist is 90° ($\phi(L) = \pm\pi/2$) and for a specified $|n'|L$ the twisting function is independent of χ, though E is not.

ACKNOWLEDGMENTS

My thanks are due to Professor Wonyong Lee and Dr. Bruce Knapp for suggesting this problem, and to Dr. Knapp for carrying out the computation described in Sec. XIV.

*This research was supported in part by the U. S. Atomic Energy Commission.

[1]H. Uberall, Phys. Rev. 103, 1055 (1956).
[2]W. K. H. Panofsky and A. N. S. Saxena, Phys. Rev. Lett. 2, 219 (1959); O. R. Frisch and D. H. Olson, Phys. Rev. Lett. 3, 141 (1959).
[3]N. Cabibbo, G. Da Prato, G. De Franceschi, and U. Mosco, Nuovo Cimento 27, 979 (1963).