LBL-4102
c.1

FAR INFRARED SPECTROSCOPY OF SOLIDS:
I. IMPURITY STATES IN Al_2O_3
II. ELECTRON-HOLE DROPLETS IN Ge

Richard Lawrence Aurbach
(Ph. D. thesis)

September 12, 1975

Prepared for the U. S. Energy Research and
Development Administration under Contract W-7405-ENG-48

For Reference

Not to be taken from this room

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY

LBL-4102
c.1

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

Table of Contents

ABSTRACT vii

I. INTRODUCTION 1

References 6

II. FOURIER TRANSFORM SPECTROMETERS 7

A. Conventional Michelson Interferometer 7

B. The Free-Standing Wire Grid Polarizer 12

C. Polarizing Michelson Interferometer 18

D. Comparison of Interferometers 30

E. Alignment 34

F. Construction Details 42

References 51

III. INSTRUMENTATION FOR FOURIER SPECTROSCOPY 52

A. Data Acquisition System 53

B. Peripherals and Interfacing 61

C. Computer Software - FTS-11 67

References 75

IV. FAR INFRARED EPR OF $Al_2O_3:Mn^{3+}$ 76

A. Experimental Details 79

B. Experimental Data 81

C. Theory 91

D. Comparison with Experiment 100

References 106

0 0 0 4 3 0 7 9 9 5

V. γ-IRRADIATED RUBY 108

A. Short Ruby 109

B. Long Ruby 111

References 115

VI. ELECTRON-HOLE DROPS IN Ge 116

A. Alpha Drop Measurements 117

B. Mie Theory 123

C. Interpretation of Alpha-Drop Results 131

D. Gamma Drop Experiments 137

E. Gamma Drop Interpretation 142

F. Magneto-Plasma Effects in Gamma Drops 151

G. Future Experiments 155

References 157

ACKNOWLEDGEMENTS 161

APPENDICES 163

A. Computer I/O Register Assignments 163

1. Teletype, KL11 164

2. Paper tape Reader/Punch, PC11 169

3. Line Frequency Clock, KW11-L 174

4. Programmable Real-Time Clock, KW11-P 176

5. Extended Arithmetic Element (EAE), KE11-A 179

6. Data Acquisition System 180

7. Oscilloscope 184

8. Card Punch 188

9. x-y Plotter 191

B. Proposed Scope Controller 195

Control Circuitry 198

x and y Axis Display Registers 202

C. Command Vocabulary 207

D. Data Analysis Algorithms 213

Fast Fourier Transform 213

Phase Correction 216

Coadding 219

Special Functions 220

References 221

0 0 0 4 3 0 7 9 9 6

FAR INFRARED SPECTROSCOPY OF SOLIDS:
I. IMPURITY STATES IN Al_2O_3
II. ELECTRON-HOLE DROPLETS IN Ge

Richard Lawrence Aurbach

Inorganic Materials Research Division, Lawrence Berkeley Laboratory
and Department of Physics; University of California
Berkeley, California 94720

ABSTRACT

Far infrared Fourier transform spectroscopy has been used to study the low lying vibronic states of Mn^{3+} in Al_2O_3 and the plasma absorption of electron-hole droplets in Ge. The transmission of Mn-doped samples of Al_2O_3 was measured in the frequency range from 3-30 cm^{-1} in applied magnetic fields up to 50 kG. Absorption lines were observed due to both ground and excited state transitions. Polarization measurements established that these absorption lines were due to electric dipole transitions. Temperature dependence measurements were used to derive a level diagram for the low lying states of Mn^{3+}. A phenomenological model based on an electronic Hamiltonian was developed which successfully describes the data. The empirically determined trigonal field and spin-orbit quenching parameters of this model are 0.7 and 0.1 respectively. This quenching is attributed to the dynamic Jahn-Teller interaction.

The plasma absorption of small (α) electron-hole drops in Ge was measured in the frequency range from 30-300 cm^{-1}. The observed absorption is in good agreement with measurements by Vavilov and other workers. A theoretical model which includes both intraband and

0 0 0 4 3 0 7 9 9 7

interband contributions to the dielectric constant in the Rayleigh limit of Mie theory is used to describe the observed lineshape. Measurements of plasma absorption of large (γ) drops in inhomogeneously stressed Ge were made in magnetic fields up to 50 kG. The lineshape at zero applied field was calculated in the large sphere limit of Mie theory including intraband terms and a zero-strain interband term. Qualitative agreement with experiment was obtained. The peak absorption shifted quadratically with applied magnetic field and the total plasma absorption increased. No oscillatory structure was observed in the field-dependence of the total absorption.

I. INTRODUCTION

In many solids the energies of interesting electronic and vibrational excitations are in the meV range. Often these excitations interact with electromagnetic radiation and may then be studied by observing the optical properties of the solid in the far infrared region of the spectrum. In this paper I discuss experiments performed on two such systems: transition metal ions in Al_2O_3 and electron-hole droplets in Ge.

Although the electronic properties of transition metal ions in solids have been extensively studied[1-4] many interesting problems remain. One area of continuing interest is systems in which the electronic excitations of isolated transition metal ions interact with the vibrational excitations of the host solid through the Jahn-Teller interaction. Examples of such systems are the $3d^4$ ions Mn^{3+} and Cr^{2+} in Al_2O_3. The Jahn-Teller effect can be understood in the following way. When an impurity ion is placed in a host crystal, the free-ion electronic energy structure is modified by local electric fields at the ion site. These electric fields arise from the local environment of the host lattice, with the largest contribution coming from the nearest neighbor atoms. If the ground state of the impurity in the host is orbitally degenerate, then it has been shown[5] that the system is unstable with respect to displacements which lower the symmetry of the environment. That is, a distortion of the local environment will split the orbitally degenerate ground state of the

0 0 0 0 4 3 0 7 9 9 8

impurity in such a way that the total energy of the system will be lowered. The distortion may be either static or dynamic. In a static distortion the time-averaged position of the atoms is changed. In some materials the static Jahn-Teller interaction is the cause of a structural phase transition. In a dynamic distortion the mean positions of the atoms remain unchanged but the atoms vibrate between lower symmetry configurations. In this case the Jahn-Teller interaction may be thought of as an interaction between the electronic states of the impurity atom and the vibrational states of the host lattice. The resultant mixed or "vibronic" states are interesting to study because they provide information on the nature of the interaction.

The transmission of several samples of Mn-doped Al_2O_3 was studied in the frequency range from 3-30 cm^{-1} in applied magnetic fields up to 50 kG. Temperature dependence measurements were used to differentiate between ground and excited state transitions. Polarization measurements revealed that the observed absorption lines were due to electric dipole transitions. Since a theory of the vibronic states of this ion is quite complicated, a phenomenological theory was developed. The electronic Hamiltonian of Mn^{3+} in the absence of the Jahn-Teller effect may be easily written. I attempted to study how well such a Hamiltonian could be used to characterize the vibronic states. By allowing the strengths of the various terms in the Hamiltonian to be adjustable parameters, it was found that this phenomenological Hamiltonian provided an excellent description of the data. The amount that various parameters were adjusted must then be related in some way to the

Jahn-Teller interaction in this system.

Another interesting system which I have studied is the electron-hole droplet in Ge. If Ge is illuminated with band-gap radiation, excitons are formed. At low temperatures and sufficiently high levels of radiation, the excitons will condense and form a liquid phase consisting of an electron-hole plasma. We may understand the condensation by following the theoretical model of Benoit a la Guillaume, et al.[6] At low excitation levels, the low-energy electronic excitations in Ge may be described as free Wannier excitons. At high light levels, however, the density of excitons is so high that considerable overlap occurs and a free exciton picture is no longer valid. The system may then be considered a two-component plasma in the Hartree-Fock approximation. The mean energy of an electron-hole pair contains contributions from the mean kinetic and exchange energies of the electrons and holes, the gap energy, and the correlation energy. Coulombic terms cancel in a mean-field approximation. Since the kinetic and exchange energies have opposite signs, a minimum can occur in the average pair energy as a function of density.

The plasma frequency of this two-component plasma is given by

$$\omega_p^{\ 2} = \frac{4\pi n e^2}{m^* \varepsilon_o} ,$$

where $n=2.3\times10^{17}$ cm^{-3} is the density, $m^*=.078\ m_e$ is the effective mass, and ε_o is the static dielectric constant of Ge. Since ω_p is about 16 meV, the far infrared region of the spectrum is the

6 6 6 / 0 2 4 0 0 0 0

appropriate frequency range to study plasma effects in the optical properties of the electron-hole drop. In unstrained Ge small (α) drops are formed with radii of around 2 microns and lifetimes of around 40 μs. In inhomogeneously stressed Ge, large (γ) drops may be formed at the maximum stress points with radii of around 300 microns and lifetimes of 400 μs. We have measured the plasma absorption of both α and γ drops in the range from 10-300 cm^{-1}. The effect of applied magnetic fields on the plasma absorption of γ-drops has also been studied.

The experiments described above were performed using the techniques of Fourier transform spectroscopy. Several advances in this technique are also described in this thesis. A new spectrometer has been developed which provides substantial improvements over conventional Michelson interferometers. This instrument, the polarizing Michelson interferometer, uses a polarizer as a beam-splitting element. Since a free standing wire grid is a nearly perfect far infrared polarizer, it can also be used as a nearly perfect beamsplitter. In addition to the spectrometer, a computer-based data acquisition and analysis system has been developed which takes advantage of recent advances in electronics technology to provide on-line data acquisition, control, and Fourier analysis functions.

In Chapter II, the polarizing interferometer is discussed. The hardware and software developed for the computer based data acquisition system are described in Chapter III. Experiments on $Al_2O_3:Mn^{3+}$ and the

phenomenological theory developed to describe the results of these experiments are detailed in Chapter IV. Results of measurements of γ-irradiated ruby are briefly described in Chapter V. The measurements of plasma absorption of electron-hole droplets in Ge are discussed in Chapter VI.

0 0 0 0 4 3 0 8 0 0 0

CHAPTER I REFERENCES

1. A. Abragam and B. Bleaney, Electron Paramagnetic Resonance of Transition Metal Ions, (Clarendon Press, Oxford, 1970).
2. W. Low, Paramagnetic Resonance in Solids, Solid State Physics, F. Seitz and D. Turnbull, ed., (Academic Press, New York, 1960).
3. C. J. Ballhausen, Introduction to Ligand Field Theory, (McGraw-Hill Book Co., New York, 1962).
4. J. S. Griffith, The Theory of Transition-Metal Ions, (Cambridge University Press, London, 1961).
5. H. A. Jahn and E. Teller, Proc. Roy. Soc. A161, 220 (1937), and Proc. Roy. Soc. A164, 117 (1938).
6. C. Benoit a la Guillaume, M. Voos, and F. Salvan, Phys. Rev. B 5, 3079 (1972).

II. FOURIER TRANSFORM SPECTROMETERS

One of the most widely used instruments for far infrared Fourier Transform Spectroscopy (FTS) is the Michelson interferometer (MI). Recently, Martin and Puplett[1] have shown that the performance of a MI may be improved by using a polarizing beamsplitter. The resultant instrument, which I call a Polarizing Michelson Interferometer (PMI) has been used successfully for astrophysical measurements.[2] In this chapter I shall discuss in detail the theory of the PMI and its construction and use for laboratory applications.

A. Conventional Michelson Interferometer

In order to provide a basis for comparison, the conventional MI is shown in Fig. 1(a). The monochromatic incident beam has amplitude i_ν. The beamsplitter B has complex amplitude reflection and transmission coefficients $re^{i\theta}$ and $te^{i\phi}$ respectively. The beam which is reflected from the movable mirror M2 has a phase shift of $e^{i\Delta}$ relative to the beam which is reflected from the fixed mirror due to the path difference δ between the two arms of the interferometer.

$$\Delta = 2\pi\nu\delta \quad , \tag{1}$$

where ν is the wavenumber in cm^{-1}. There are two output beams, labeled i_{det} and i_{src}. The intensities of these beams are given by

$$\begin{aligned} I_{det} &= \left| i_\nu re^{i\theta}\ te^{i\phi}\ (1 + e^{i\Delta}) \right|^2 \\ &= 2I_\nu\ RT(1 + \cos\Delta) \end{aligned} \tag{2}$$

0 0 0 0 4 3 0 8 0 1

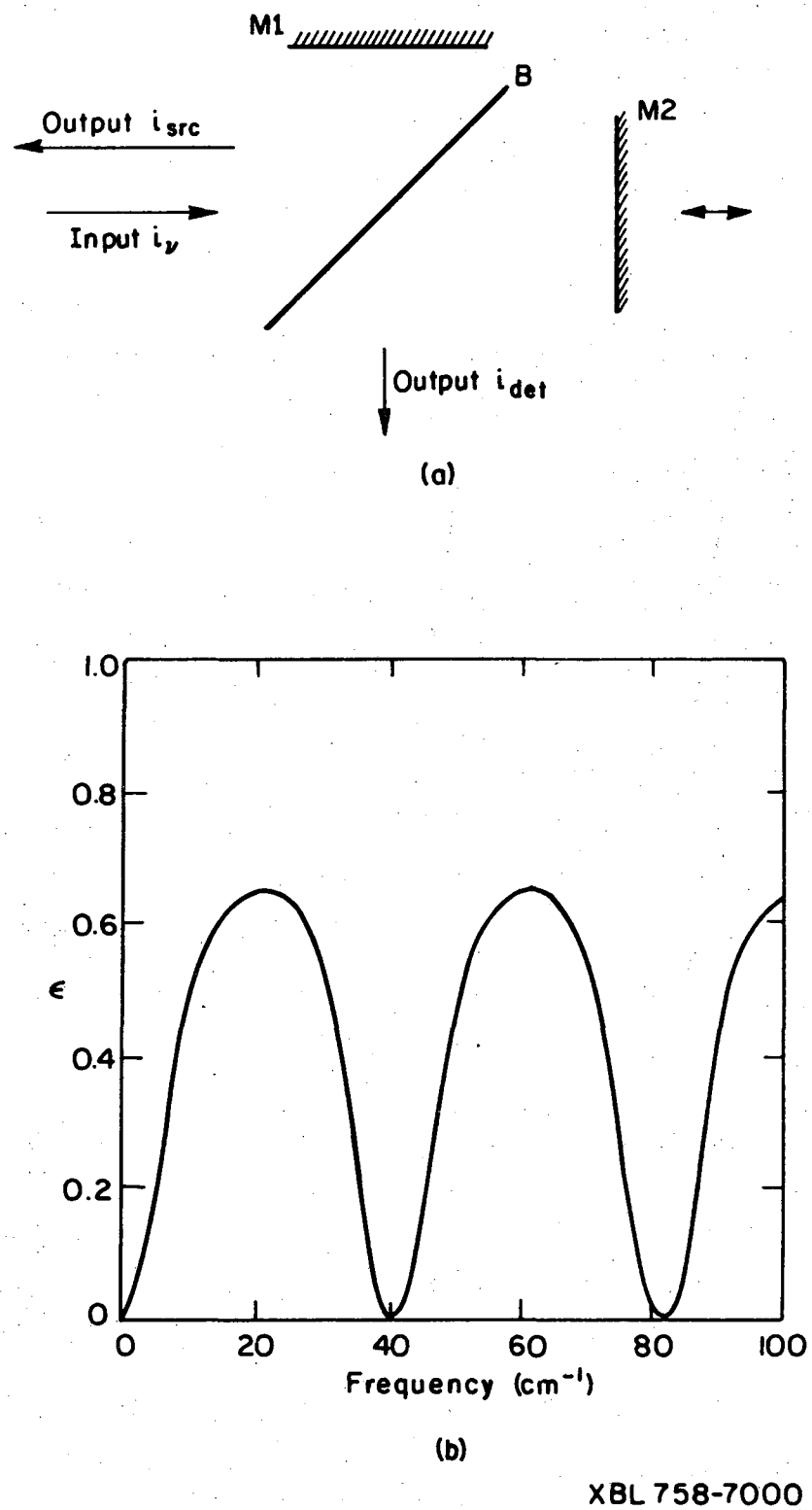


Fig. 1. Michelson Interferometer.

and

$$I_{src} = |i_\nu \, (r^2 e^{2i\theta} + t^2 e^{2i\phi} \, e^{i\Delta})|^2$$

$$= I_\nu \, \{R^2 + T^2 + 2RT \cos \, [2(\theta-\phi) + \Delta]\} \qquad (3)$$

respectively, where $I_\nu = i_\nu^2$, $R = r^2$, and $T = t^2$. If the beamsplitter is nonabsorbing, conservation of energy requires that $I_\nu = I_{det} + I_{src}$. This will be the case provided that[3]

$$\theta-\phi = (2n+1) \frac{\pi}{2}, \qquad n = 1,2,3,\ldots \quad . \qquad (4)$$

Most laboratory interferometers, such as the one shown in Fig. 2, collect only the beam I_{det}. Double beam operation is possible, however. Suppose that we collect both output beams, but with a sample of transmittance $\mathcal{T}$ in the I_{det} beam. Then $I = I_{src} + \mathcal{T} I_{det}$. Assuming that the beamsplitter is nonabsorbing (R+T=1), we get

$$I = I_\nu \{1-2RT(1-\mathcal{T}) - 2RT(1-\mathcal{T}) \cos\Delta\} \, . \qquad (5)$$

Conventional far infrared MI's use a dielectric beamsplitter made of a thin film of Mylar. If we neglect the absorption of the Mylar, the properties of such a beamsplitter may easily be calculated.[4] The transmission and reflection are determined by the single surface Fresnel coefficients $\mathcal{R}$ and $\mathcal{T}$ and the Fabry-Perot interference inside the dielectric film. For light polarized perpendicular and parallel to the plane of incidence, the Fresnel coefficients are[5]

0 0 0 4 3 0 8 0 0 2

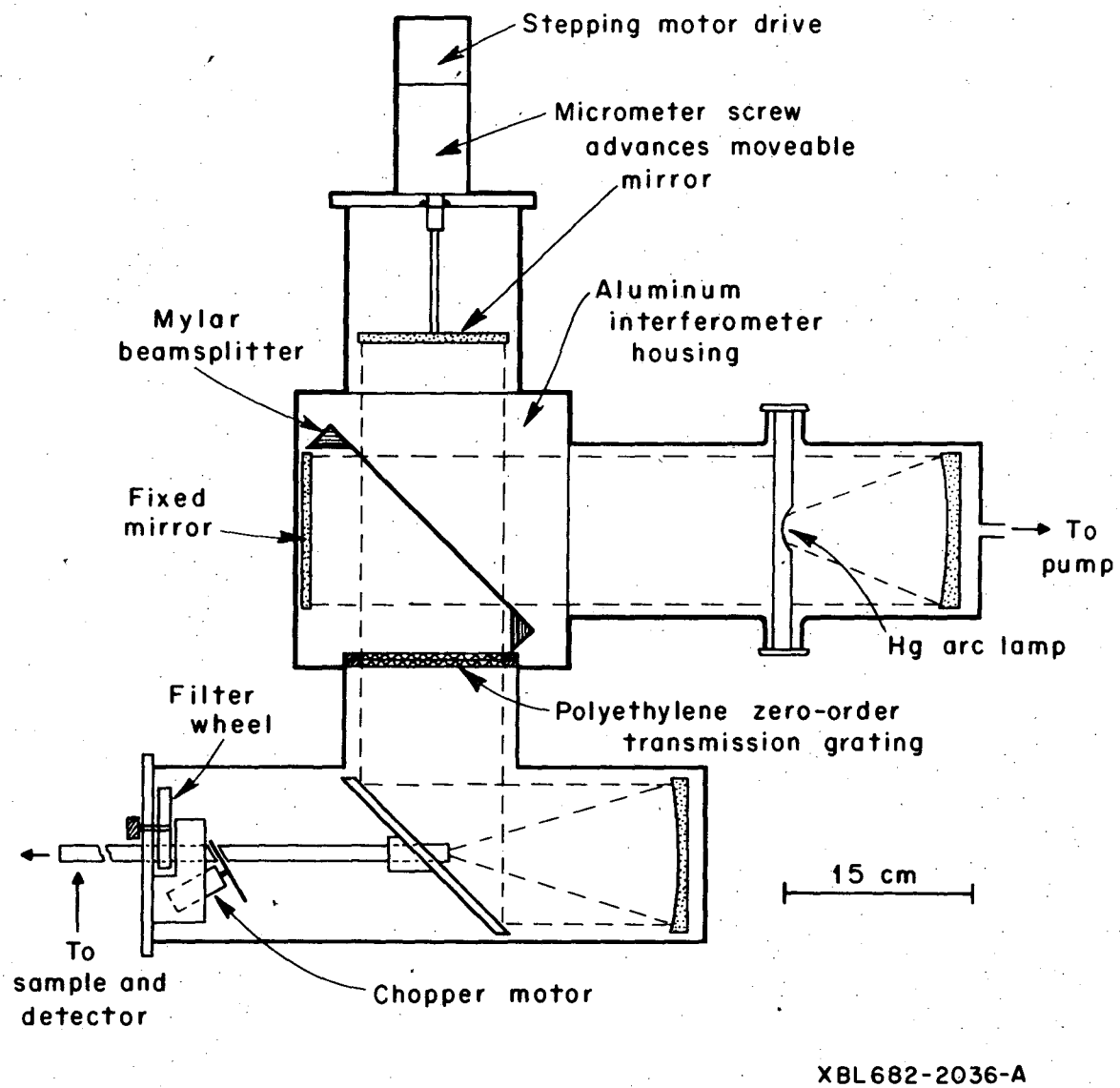


Fig. 2. Laboratory Michelson Interferometer.

$$\mathcal{R}_{\perp} = \frac{\sin^2(\theta_i - \theta_t)}{\sin^2(\theta_i + \theta_t)} \quad , \qquad \mathcal{T}_{\perp} = 1 - \mathcal{R}_{\perp} \ , \tag{6}$$

and

$$\mathcal{R}_{\parallel} = \frac{\tan^2(\theta_i - \theta_t)}{\tan^2(\theta_i + \theta_t)} \quad , \qquad \mathcal{T}_{\parallel} = 1 - \mathcal{R}_{\parallel} \ , \tag{7}$$

respectively, where θ_i is the angle of incidence (45° in this case) and θ_t is the angle of the refracted ray inside the dielectric. Fabry-Perot effects are expressed in terms of a phase function ψ.

$$\psi = 4\pi n \nu d \cos\theta_t \ , \tag{8}$$

where n is the index of refraction (1.75 for Mylar in the far infrared), ν is the wavenumber in cm^{-1} and d is the thickness of the beamsplitter. The beamsplitter efficiency for one polarization is $\varepsilon = 4RT$.

$$\varepsilon = \frac{8\mathcal{R}\mathcal{T}^2(1 - \cos\psi)}{(1 + \mathcal{R}^2 - 2\mathcal{R}\cos\psi)^2} \quad . \tag{9}$$

This function has zeros when $\psi=0, 2\pi, 4\pi, \ldots$. For frequencies near zero, ε is proportional to ν^2. The maxima of ε occur when $\psi=\pi, 3\pi, 5\pi, \ldots$, with the peak efficiency being given by

$$\varepsilon_{max} = \frac{16\mathcal{R}\mathcal{T}^2}{(1+\mathcal{R})^4} \quad . \tag{10}$$

The efficiency for unpolarized light is the average of the efficiencies for the two orthogonal polarizations

$$\varepsilon = \frac{1}{2}(\varepsilon_{\parallel} + \varepsilon_{\perp}) \ ,$$

0 0 0 4 3 0 8 0 0 3

which for Mylar has a peak value of roughly 65%. In Fig. 1(b) the theoretical efficiency of a .00762 cm (3 mil) Mylar beamsplitter is plotted.

B. The Free-Standing Wire Grid Polarizer

The major weakness of the conventional MI is the beamsplitter efficiency. The zeros in the efficiency create considerable inconvenience in making broad-band measurements. In addition the peak efficiency is not as high as one would like. These weaknesses are intrinsic in the use of a dielectric beamsplitter. A new type of beamsplitter is required if the efficiency is to be improved.

It is apparent that a perfect polarizer would be a perfect beamsplitter for linearly polarized light. At far infrared wavelengths, a one-dimensional free-standing grid of parallel wires makes a nearly ideal polarizer for wavelengths long compared to the wire spacing. Moreover, thin tungsten wires are sufficiently strong that construction of freestanding wire polarizers is feasible for use below 200 cm^{-1}. Since freestanding polarizers lack a dielectric substrate, they are particularly appealing in this context.

The properties of a wire grid polarizer are most easily discussed in terms of the microwave equivalent circuits[6] describing the polarizer. We consider a one-dimensional array of wires of diameter d with a center-to-center spacing a (Fig. 3(a)) oriented at an angle θ with respect to the incident radiation. There are two cases. The electric field may be oriented either parallel to the wires ("inductive grid", Fig. 3(b)) or perpendicular to the wires ("capacitive grid", Fig. 3(c)).

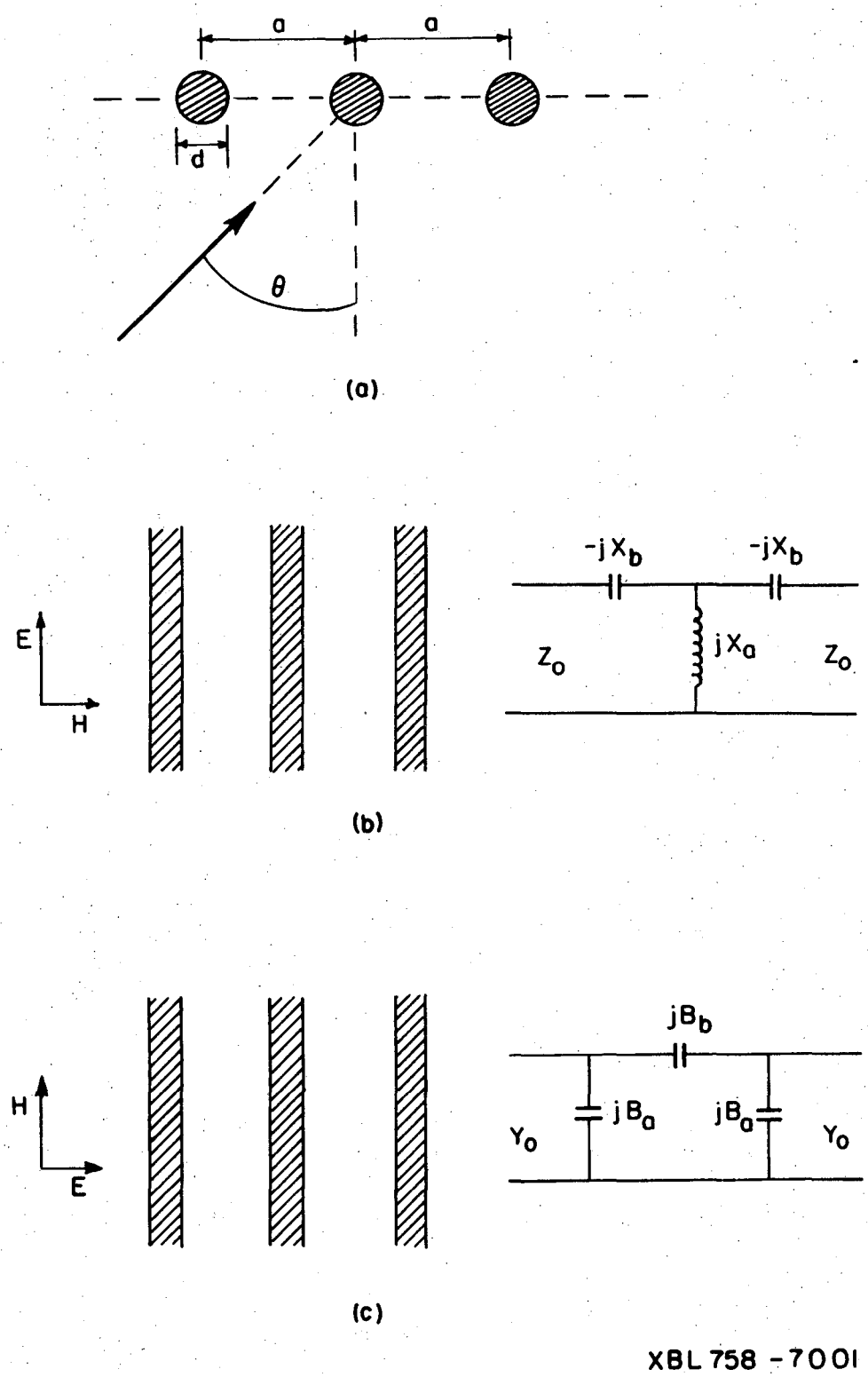


Fig. 3. Wire Grid Polarizer--Equivalent Circuits.

0 0 0 0 4 3 0 8 0 0 4

For the inductive grid, the polarizer is described by a parallel inductor. The finite wire size is included by adding the series capacitors as is shown in the equivalent circuit in Fig. 3(b). The impedance Z of this circuit is computed by terminating the circuit with Z_o, the impedance of free space.

$$Z = \frac{2X_aX_b - X_b^2 + jZ_o(X_a - X_b)}{Z_o + j(X_a - X_b)} \tag{11}$$

The amplitude reflectivity and transmittivity coefficients Γ_i and τ_i are given by[7]

$$\Gamma_i = \frac{Z - Z_o}{Z + Z_o} \quad , \qquad \tau_i = 1 + \Gamma_i \ . \tag{12}$$

$$\Gamma_i = \frac{2X_aX_b - X_b^2 - Z_o^2}{2X_aX_b - X_b^2 + Z_o^2 + 2jZ_o(X_a - X_b)} \quad . \tag{13}$$

The energy reflectivity and transmittivity coefficients are $|\Gamma_i|^2$ and $|\tau_i|^2$ respectively. The equivalent circuit is valid for

$$a(1 + \sin\theta)/\lambda < 1. \tag{14}$$

Marcuvitz[8] has given approximate expressions for the reactances X_a and X_b.

$$\frac{X_a}{Z_o} = \frac{a\cos\theta}{\lambda}\left\{\ln\frac{a}{2\pi d} + \frac{1}{2}\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty}\left[\left(m^2 - \left(\frac{a\cos\theta}{\lambda}\right)^2 + \frac{2ma\sin\theta}{\lambda}\right)^{-\frac{1}{2}} - \frac{1}{|m|}\right]\right\}. \tag{15a}$$

$$\frac{X_b}{Z_o} = \frac{a\cos\theta}{\lambda}\left(\frac{2\pi d}{a}\right)^2 \quad . \tag{15b}$$

The computed energy reflectivity $|\Gamma_i|^2$ has been plotted in Fig. 4(a) for the case of d = .00254 cm, a = .00762 cm (1 mil wire with 3 mil spacing). In the dc limit ($\lambda \to \infty$, $\nu \to o$) the grid is perfectly reflecting. The reflectivity falls slowly for finite frequencies until $a(1 + \sin\theta)$ becomes comparable with λ. At this point the reflectivity falls precipitously.

For the capacitive grid, a circuit description in terms of admittances is more useful than one in terms of impedances. The grid is described by a parallel capacitance, with finite wire size effects adding a series capacitance as is shown in Fig. 3(c). The effective admittance of the circuit is

$$Y = \frac{-B_a(2B_a+B_b) + jY_o(B_a+B_b)}{Y_o + j(B_a+B_b)} \quad . \tag{16}$$

The amplitude reflectivity coefficient is given by

$$\Gamma_c = \frac{Y_o - Y}{Y_o + Y} \quad , \tag{17}$$

with

$$\Gamma_c = \frac{Y_o^2 + B_a^2 + 2B_aB_b}{Y_o^2 - B_a^2 - 2B_aB_b + 2jY_o(B_a+B_b)} \quad . \tag{17}$$

Once again the validity of this description is limited by Eq. (14). Marcuvitz solved the capacitive grid problem only for the case of normal incidence. His solutions may be extended approximately to arbitrary angle of incidence by noticing that in the inductive grid case (Eq. (15)) most of the angular dependence is accounted for by

replacing λ by $\lambda/\cos\theta$ in the leading factors for the reactances. A similar substitution for the capacitive grid case gives approximate expressions[9] for B_a and B_b.

$$\frac{B_a}{Y_o} = \frac{2a\cos\theta}{\lambda}\left(\frac{\pi d}{2a}\right)^2 \frac{1}{A_2} , \tag{18a}$$

$$\frac{B_b}{Y_o} = \frac{\lambda}{2a\cos\theta}\left(\frac{2a}{\pi d}\right)^2 A_1 - \frac{a\cos\theta}{\lambda}\left(\frac{\pi d}{2a}\right)^2 \frac{1}{A_2} , \tag{18b}$$

where A_1 and A_2 are given by

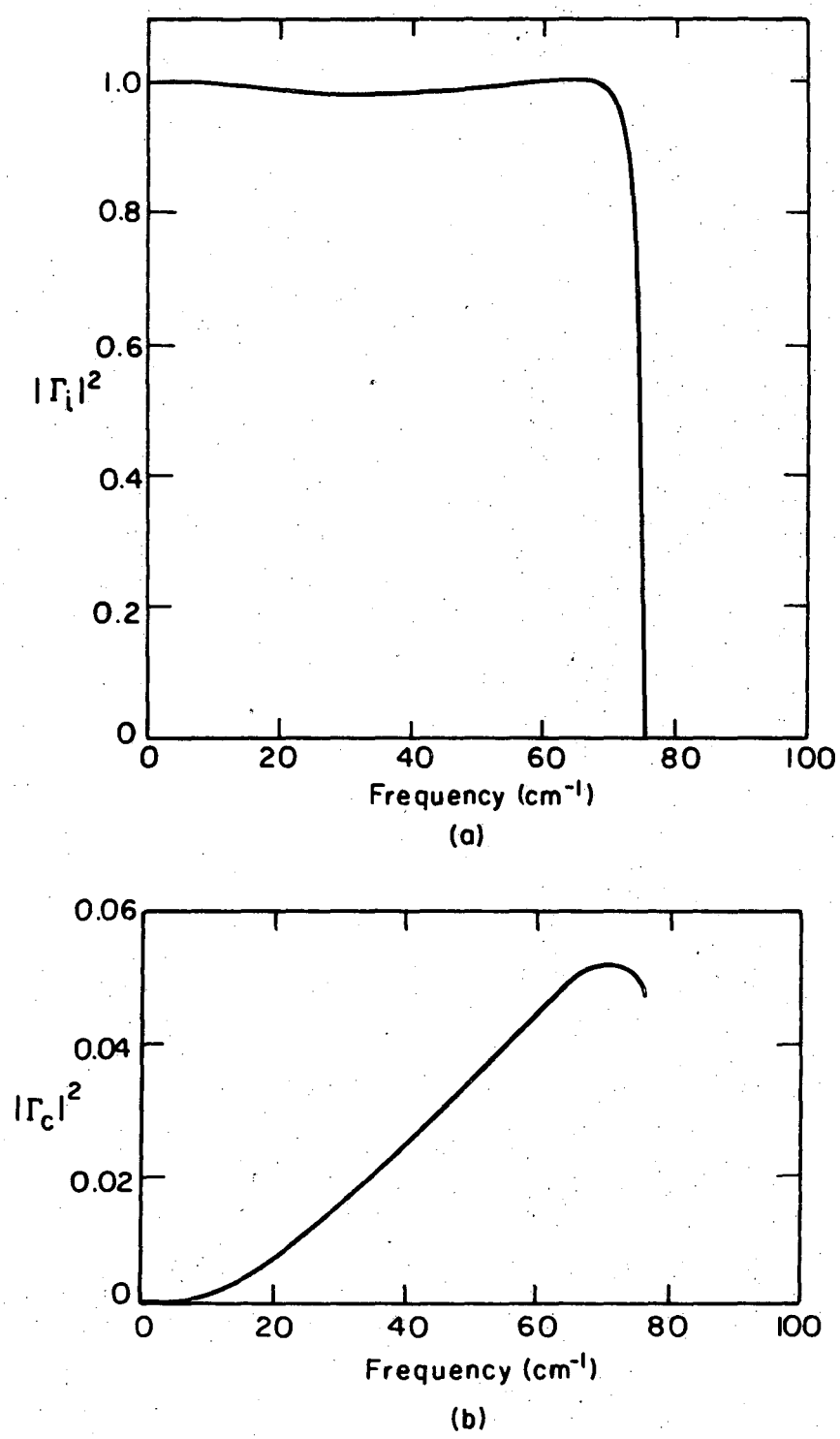
$$A_1 = 1 + \frac{1}{2}\left(\frac{\pi d}{\lambda}\right)^2\left(\ln\frac{a}{\pi d} + \frac{3}{4}\right) + \left(\frac{\pi d}{\lambda}\right)^2 \sum_{m=1}^{\infty} \cos^2\frac{m\pi}{2}\left[\frac{1}{\sqrt{m^2 - \left(\frac{2a}{\lambda}\right)^2}} - \frac{1}{m}\right] , \tag{19a}$$

and

$$A_2 = 1 + \frac{1}{2}\left(\frac{\pi d}{\lambda}\right)^2\left(\frac{11}{4} - \ln\frac{a}{\pi d}\right) + \frac{1}{6}\left(\frac{\pi d}{2a}\right)^2 - 2\left(\frac{\pi d}{2a}\right)^2 \sum_{m=1}^{\infty} \cos^2\frac{m\pi}{2}\left[m - \frac{1}{2m}\left(\frac{2a}{\lambda}\right)^2 - \sqrt{m^2 - \left(\frac{2a}{\lambda}\right)^2}\right] . \tag{19b}$$

The energy reflectivity $|\Gamma_c|^2$ has been plotted in Fig. 4(b) for the case of d=.00254 cm, a=.00762 cm. The capacitive grid is perfectly transparent in the dc limit.

In these expressions, beamsplitter absorption has been neglected. Casey and Lewis[10] have shown that for the inductive grid the power absorbed is



XBL 758-7002

Fig. 4. Wire Grid Polarizer--Reflectivities.

0 0 0 4 3 0 8 0 0 5

$$A_i = \frac{a}{2\pi d} \left(\frac{c \cos\theta}{\sigma\lambda}\right)^{1/2} |\Gamma_i|^2 ,$$

where σ is the conductivity of the wires and c is the velocity of light. For most cases A_i is less than one per cent. Since in the capacitive grid, induced currents flow across rather than along the wires, the absorbance is expected to be even less. Therefore the approximation of neglecting beamsplitter absorption is a good one.

C. Polarizing Michelson Interferometer

We may now study the properties of an interferometer which uses the polarizer described above as a beamsplitter. For this discussion we will refer all polarization vectors to the beamsplitter polarization directions $\hat{i}$ ($\underset{\sim}{E} \parallel$ wires) and $\hat{c}$ ($E \perp$ wires). A schematic diagram of the PMI is shown in Fig. 5. For convenience we will assume that the input polarizer P_1 is fixed at an angle of 45° with respect to the beamsplitter B and that P_2 rotates at angular frequency ω. In the actual device which has been constructed, P_1 was rotated while P_2 was held fixed. There is no conceptual difference between these two cases--the analysis is simpler however if the time-dependence of the polarization state enters at the end of the calculation.

We assume that unpolarized radiation of amplitude $\sqrt{I}$ is incident on P_1. The polarizer transmits $\frac{\sqrt{I}}{2}(\hat{i} + \hat{c})$. The beamsplitter reflectivities and transmittivities for $\hat{i}$ and $\hat{c}$ polarized radiation are Γ_i, τ_i, Γ_c, and τ_c, respectively. Therefore the light reflected by the beamsplitter toward mirror M_1 is $\frac{\sqrt{I}}{2}(\Gamma_i\hat{i} + \Gamma_c\hat{c})$ while the

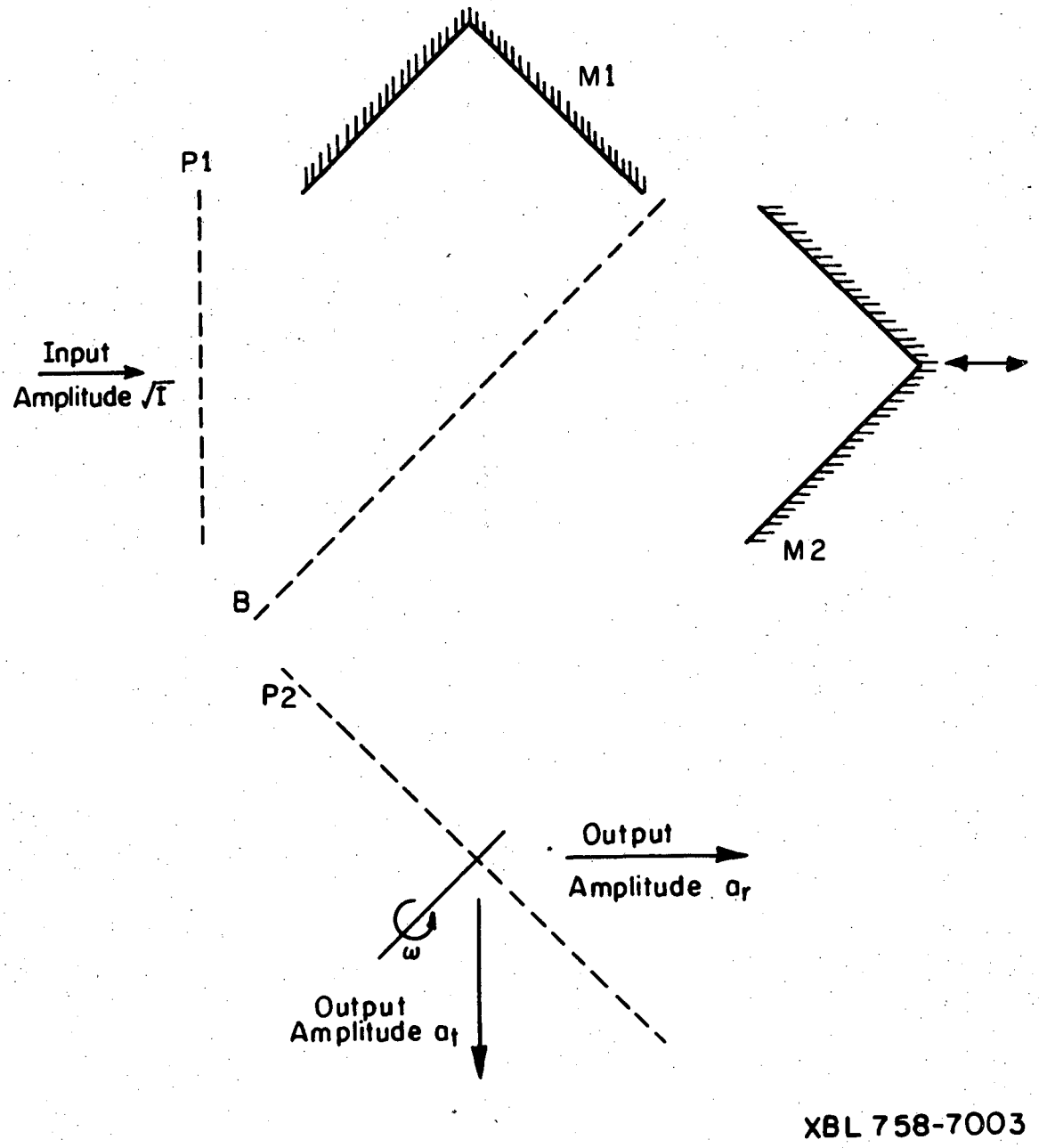


Fig. 5. Polarizing Michelson Interferometer--
Schematic Diagram.

0 0 0 4 3 0 8 0 0 6

radiation transmitted toward mirror M_2 is $\frac{\sqrt{I}}{2}$ $(\tau_i\hat{i} + \tau_c\hat{c})$. The mirrors M_1 and M_2 are dihedral mirrors oriented at 45° with respect to B. They rotate the plane of polarization by 90°. These mirrors will be discussed in more detail below. The beam reflected from M_1 is therefore $\frac{\sqrt{I}}{2}$ $(\Gamma_i\hat{c} - \Gamma_c\hat{i})$. The beam reflected from M_2, in addition to the rotation of the plane of polarization also experiences a phase shift Δ given by Eq. (1) due to the difference in optical path lengths in the two arms of the interferometer. Therefore this beam is $\frac{\sqrt{I}}{2}$ $(\tau_i\hat{c} - \tau_c\hat{i})e^{i\Delta}$. The part of the beam from M_1 which is transmitted is $\frac{\sqrt{I}}{2}$ $(\Gamma_i\tau_c\hat{c} - \Gamma_c\tau_i\hat{i})$ while the part of the beam from M_2 which is reflected is $\frac{\sqrt{I}}{2}$ $(\tau_i\Gamma_c\hat{c} - \tau_c\Gamma_i\hat{i})e^{i\Delta}$. At polarizer P_2 the total amplitude is

$$\frac{\sqrt{I}}{2}\left\{\left(\Gamma_i\tau_c + \tau_i\Gamma_c e^{i\Delta}\right)\hat{c} - \left(\Gamma_c\tau_i + \tau_c\Gamma_i e^{i\Delta}\right)\hat{i}\right\} \quad . \tag{20}$$

In general the radiation is elliptically polarized. The rotating polarizer P_2 transmits $\frac{1}{\sqrt{2}}$ $\{\hat{i}(\cos\omega t + \sin\omega t) + \hat{c}(\cos\omega t - \sin\omega t)\}$ and reflects $\frac{1}{\sqrt{2}}$ $\{\hat{i}(\cos\omega t - \sin\omega t) - \hat{c}(\cos\omega t - \sin\omega t)\}$, where we have chosen the zero of phase for P_2 oriented at 45° with respect to the beamsplitter. The transmitted amplitude is

$$\begin{aligned} a_t = \frac{\sqrt{I}}{2\sqrt{2}} \; & \{-(\Gamma_c\tau_i + \tau_c\Gamma_i e^{i\Delta})(\cos\omega t + \sin\omega t) \\ & + (\Gamma_i\tau_c + \tau_i\Gamma_c e^{i\Delta})(\cos\omega t - \sin\omega t) \quad , \end{aligned} \tag{21a}$$

and the amplitude reflected is

$$a_r = \frac{-\sqrt{I}}{2\sqrt{2}} \{(\Gamma_c\tau_i + \tau_c\Gamma_i e^{i\Delta})(\cos\omega t - \sin\omega t)$$

$$+ (\Gamma_i\tau_c + \tau_i\Gamma_c e^{i\Delta})(\cos\omega t + \sin\omega t)\} \ . \tag{21b}$$

For a perfect beamsplitter ($|\Gamma_i| = |\tau_c| = 1$, $|\Gamma_c| = |\tau_i| = 0$), the expressions simplify greatly and the transmitted and reflected intensities are

$$I_t^o = \frac{I}{4}(1 + \cos\Delta \cos 2\omega t) \ , \tag{22a}$$

$$I_r^o = \frac{I}{4}(1 - \cos\Delta \cos 2\omega t) \ . \tag{22b}$$

Detecting synchronously at 2ω, the transmitted and reflected intensities are 180° out of phase and of magnitude $\frac{I}{2}\cos\Delta$. If the beamsplitter is not perfect, the expressions for the intensities are much more complicated. Writing $\Gamma_i = |\Gamma_i|e^{i\phi_{\Gamma i}}$ and similarly for Γ_c, τ_i, and τ_c, we get

$$I_{t,r} = \frac{I}{4}\{|\Gamma_c|^2|\tau_i|^2 + |\Gamma_i|^2|\tau_c|^2 + 2|\Gamma_i\Gamma_c\tau_i\tau_c|\cos\delta\phi\cos\Delta$$

$$\pm [|\Gamma_2|^2|\tau_i|^2 + |\Gamma_i|^2|\tau_c|^2]\cos\Delta\cos 2\omega t$$

$$\pm 2|\Gamma_i\Gamma_c\tau_i\tau_c|\cos\delta\phi\cos 2\omega t\} \ , \tag{23}$$

where in the second and third terms, the plus signs refer to the transmitted intensity and the minus signs refer to the reflected intensity. The phase $\delta\phi$ is

$$\delta\phi = \phi_{\Gamma i} - \phi_{\Gamma c} - \phi_{\tau i} + \phi_{\tau c} \ . \tag{24}$$

0 0 0 0 4 3 0 8 0 0 7

Since $\tau = 1 + \Gamma$ and $|\Gamma|^2 + |\tau^2| = 1$, it is easy to show that the phases of Γ and τ differ by $\frac{\pi}{2}$. Hence $\delta\phi=0$. The transmitted and reflected intensities contain a dc term and two terms at 2ω. Synchronously detecting at 2ω, we measure

$$I_{t,r} = \frac{I}{2} \{\mp[|\Gamma_c\tau_i|^2 + |\Gamma_i\tau_c|^2]\cos\Delta$$

$$\mp 2|\Gamma_i\Gamma_c\tau_i\tau_c|\}. \tag{25}$$

Comparing these expressions with the results for a perfect polarizing beamsplitter, we see that a beamsplitter efficiency ε may be defined by

$$\varepsilon = |\Gamma_c\tau_i|^2 + |\Gamma_i\tau_c|^2 \quad . \tag{26}$$

Non-ideality of the beamsplitter also leads to a small offset term.

The properties of the beamsplitter are characterized by four independent quantities, the real and imaginary parts of Γ_i and Γ_c. The real and imaginary parts of τ_i and τ_c are determined by the relations $\tau_i = 1 + \Gamma_i$ and $\tau_c = 1 + \Gamma_c$. The combinations of these parameters which are useful for the characterization of a beamsplitter are ε, $a = 2|\Gamma_i\Gamma_c\tau_i\tau_c|$, and the fractional polarization P. The fractional polarization is defined as

$$P = \frac{|\Gamma_i|^2 - |\Gamma_c|^2}{|\Gamma_i|^2 + |\Gamma_c|^2} \quad . \tag{27}$$

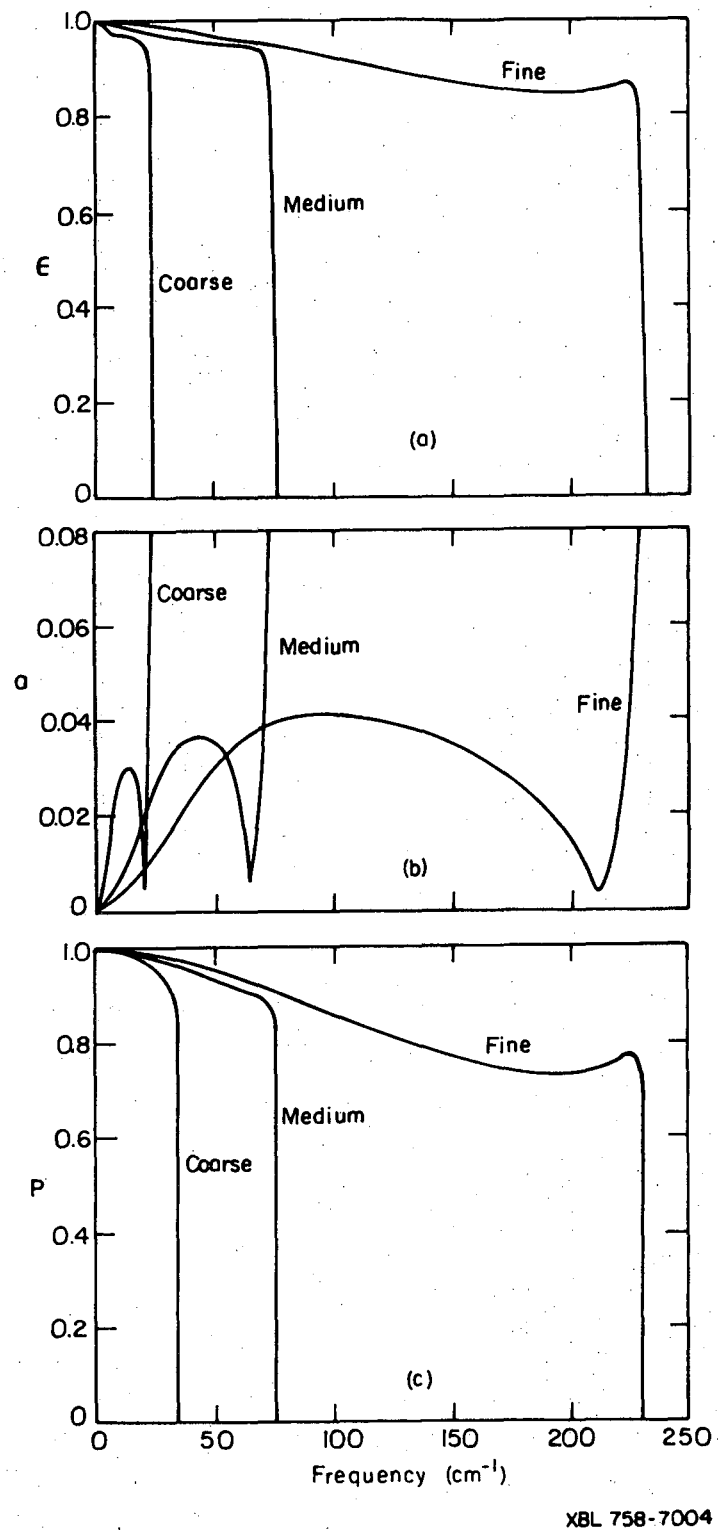


Fig. 6. Beamsplitter Properties of Wire Grid Polarizers.

0 0 0 4 3 0 8 0 0 8

Table I. Beamsplitter Parameters

Beamsplitter	a(cm)	d(cm)	$v_{max}(cm^{-1})$
Fine	.00254	.00127	230
Medium	.00762	.00254	77
Coarse	.0254	.00508	23

In Fig. 6 we have plotted these three quantities for fine, medium, and coarse beamsplitters. Table 1 gives the spacing and diameters of the wires for these three cases. The parameter ν_{max} is the frequency at which the equivalent circuit formalism breaks down. It is given by $a(1 + \sin\theta)\nu_{max} = 1$.

As the spacing and diameter of the wires is reduced in order to increase ν_{max}, a point will be reached when the beamsplitter can no longer be made free-standing. It is reasonable to ask, therefore, what effect a substate would have on beamsplitter performance. This problem is exceedingly complicated because it is important to include multiple reflections both within the dielectric layer and between the grid and the dielectric surfaces. An approximate calculation has been done for a perfect polarizer imbedded in the center of a dielectric. This approximation is illustrated in Fig. 7. For $\hat{i}$ polarized radiation, a perfect polarizer is a perfect reflector. The reflected wave, then, includes all of the multiple reflections involving the front surface of the dielectric and the polarizer, which is located half-way into the film of thickness h. (Fig. 7(a)) Assuming a dielectric index n and an angle θ_t for the refracted ray in the dielectric, we define $\delta = 2\pi\nu nh\cos\theta_t$. The amplitude reflection coefficient for the composite beamsplitter is

$$\Gamma_i = \frac{r + e^{i(\delta+\delta_m)}}{1 + re^{i(\delta+\delta_m)}} , \tag{28}$$

0 0 0 0 4 3 0 8 0 0 9

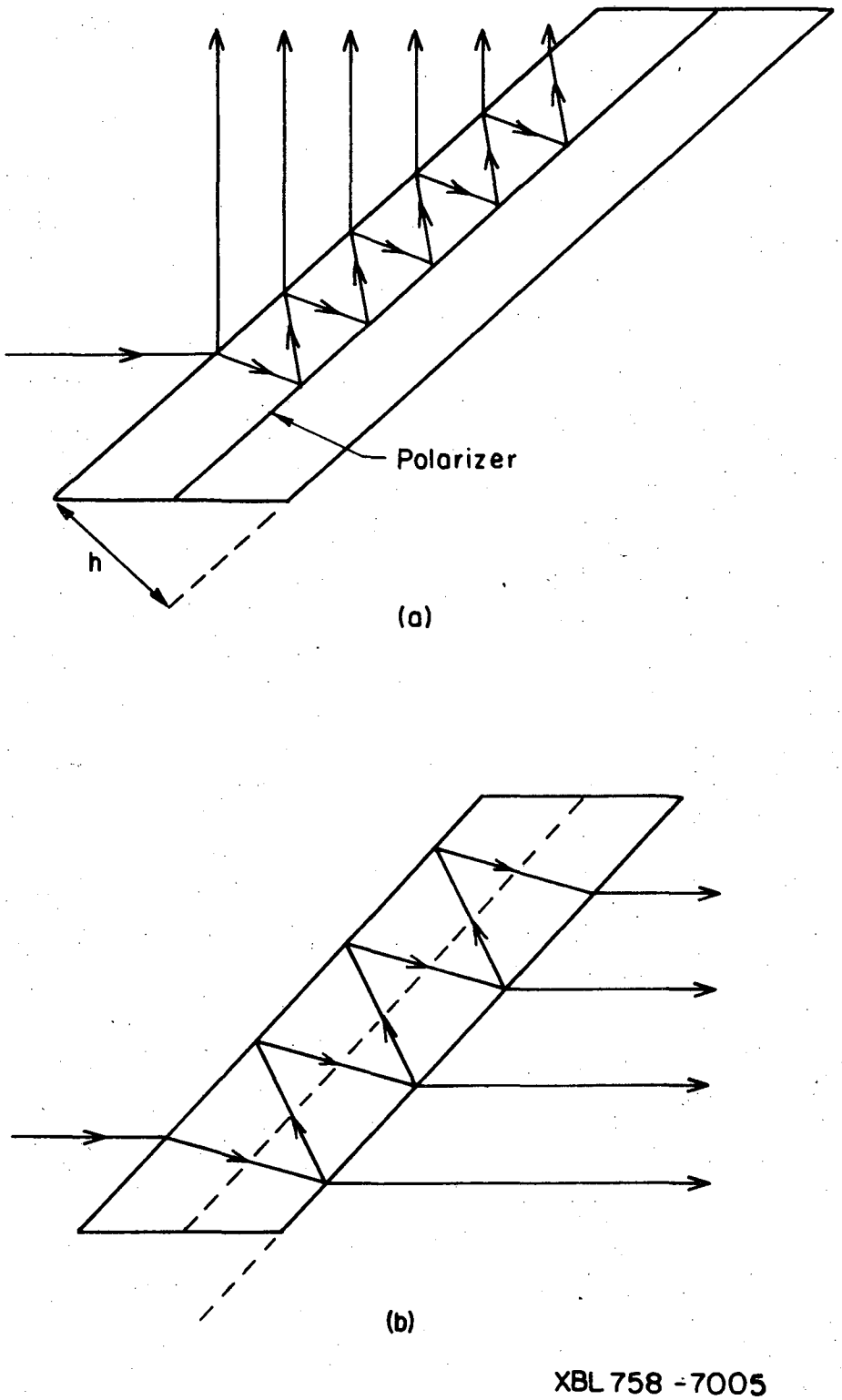


Fig. 7. Composite Beamsplitter Model.

where r is the amplitude reflectivity from the front surface of the dielectric and δ_m is the phase shift upon metallic reflection. For $\hat{c}$ polarized radiation, a perfect polarizer is perfectly transparent. The amplitude transmission coefficient is given by the usual Fabry-Perot result for a dielectric layer of thickness h.

$$\tau_c = \frac{Te^{i\delta}}{1-Re^{2i\delta}} \tag{29}$$

where R and T are the energy reflection and transmission coefficients for the dielectric interface. By assumption, $\tau_i = \Gamma_c = 0$. The output at 2ω of a PMI with a composite beamsplitter is given by Eq. (25). Substituting the above expressions in (25) and noticing that $|\Gamma_i|^2 = 1$, we get

$$I = \frac{I_o}{2}\left(\frac{T^2}{1+R^2-2R\cos\delta}\right)\cos\Delta \ . \tag{30}$$

The efficiency of the PMI is given by the Fabry-Perot transmission function and therefore oscillates as a function of frequency. For a Mylar substrate, the efficiency oscillates between 70 and 100% as a function of frequency below the polarizer cutoff. To the limit of validity of this crude approximation, a composite beamsplitter is useful. Since a better approximation is difficult, this result should be checked experimentally.

In a real instrument, two comments about the polarizers P_1 and P_2 must be made. First, these polarizers must be tilted relative to the propagation directions in the interferometer in order to eliminate the

0 0 0 4 3 0 8 0 1 0

effect of multiple passes of radiation through the instrument. There is a substantial amount of reflected power from polarizer P_2 (Eq. (21b)) which would otherwise pass back through the interferometer, reflect from P_1 and pass through again. These multiple passes would lead to erroneous signals. Also consider the recombination of radiation from the corner mirrors at the beamsplitter. A real beamsplitter reflects some of the radiation from M_1 and transmitts some of the radiation from M_2. This radiation will strike P_1. If P_1 is not tilted, it can reflect from P_1 and pass through the interferometer again. Therefore it is essential that P_1 be inclined so that reflections from it are removed from the beam. It is desirable to tilt P_2 as well so that there are no standing waves set up in the output of the interferometer.

Secondly we have assumed so far that the polarizers P_1 and P_2 are perfect. This is clearly not the case. If we assume a perfect beamsplitter we may calculate the effect of two identical imperfect polarizers characterized by the parameters τ_c, τ_i, Γ_c, and Γ_i. The input polarizer transmits a main beam of amplitude $\frac{\sqrt{I}}{2}\tau_c(\hat{i}+\hat{c})$ and a leakage beam of amplitude $\frac{\sqrt{I}}{2}\tau_i(\hat{i}-\hat{c})$. Since we have assumed that the incident light is unpolarized, these two beams are uncorrelated. Therefore we treat them separately and add their output intensities. When the output polarizer is oriented perpendicular to the input polarizer, its transmission relative to the beamsplitter coordinate vectors is $\frac{1}{\sqrt{2}}[(\tau_c-\tau_i)\hat{i} - (\tau_c+\tau_i)\hat{c}]$. We define the phases of the complex quantities $\tau_c+\tau_i$ and $\tau_c-\tau_i$ by δ_+ and δ_-. If $|\tau_c| \gg |\tau_i|$ (which will be the case except near the cutoff frequency ν_{max} of the

polarizer) $\delta_+ \approx \delta_- \approx \phi_{\tau c}$, the phase of τ_c. Using these phase functions we calculate the intensity transmitted by P_2 due to the main and leakage beams. Adding intensities we get at 2ω

$$I = \frac{I}{2} (\tau_c^2 - \tau_i^2)^2 \cos(\delta_+ - \delta_-) \cos\Delta \ . \tag{31}$$

This enables us to define a polarizer efficiency η_T for the signal transmitted by P_2.

$$\eta_T = (\tau_c^2 - \tau_i^2)^2 \cos(\delta_+ - \delta_-) \ . \tag{32}$$

Since $\delta_+ \approx \delta_-$ for frequencies much below ν_{max}, the cosine is approximately unity. Repeating the calculation for the output beam reflected from P_2,

$$\eta_R = (\tau_c^2 - \tau_i^2)(\Gamma_i^2 - \Gamma_c^2) \cos(\gamma_+ - \gamma_-) \ , \tag{33}$$

where γ_+ and γ_- are the phases of $\Gamma_i + \Gamma_c$ and $\Gamma_i - \Gamma_c$ respectively. Since $|\Gamma|^2 + |\tau|^2 = 1$, we may rewrite (33) as

$$\eta_R = (\tau_c^2 - \tau_i^2)^2 \cos(\gamma_+ - \gamma_-) \ . \tag{34}$$

In general the cosine terms in η_T and η_R are different, so that the powers in the transmitted and reflected beams are not the same. This will lead to a small beam imbalance in a double-beam geometry.

One final comment must be made concerning a real PMI. At frequencies greater that ν_{max}, the beamsplitter acts as a diffraction grating and disperses the radiation. If this dispersed high frequency radiation remains in the plane of the interferometer, then it can strike the dihedral mirrors, recombine, etc. The result is false radiation which

0 0 0 0 4 3 0 8 0 1 1

may be modulated by the movement of M_2. This radiation can, of course, be removed by filtering the output before the detector. A better way, however, is simply to orient the beamsplitter so that the wires are parallel to the plane of the instrument. The high frequency radiation is then dispersed out of the beam, and extra filtration is not required to remove it.

D. Comparison of Interferometers

In order to compare the MI discussed in Sec. A with the PMI discussed above, we rewrite Eq. (2) in terms of the conventional beamsplitter efficiency ε_{MI},

$$I_{MI} = \frac{1}{2} I_o \varepsilon_{MI} \cos\Delta + \frac{1}{2} I_o \varepsilon_{MI} \tag{35}$$

If a MI is chopped with a two-bladed chopper rotating at frequency ω, I_{MI} will be the signal detected synchronously at 2ω. The analogous expression for the PMI is derived from Eq. (25).

$$I_{PMI} = \frac{1}{2} I_o \varepsilon_{PMI} \cos\Delta + \frac{1}{2} I_o a \ , \tag{36}$$

where a was defined above.

The most important difference in the expressions is due to the difference between the beamsplitter efficiencies. The efficiency for a typical Mylar beamsplitter is shown in Fig. 1(b). At low frequencies ε_{MI} is proportional to ν^2. The peak efficiency is about 65%. For a polarizing beamsplitter, however, the low frequency efficiency is unity and falls showly with increasing frequency up to ν_{max}. Clearly at low frequencies for which $\varepsilon_{PMI} > \varepsilon_{MI}$, the PMI is more efficient.

Also the zeros of ε_{MI} at finite frequency are a great inconvenience in measuring the properties of a sample over a broad range in frequency, since several changes of beamsplitter will be necessary in order to maintain good efficiency throughout the measurement range. Only a single polarizing beamsplitter is necessary for such a measurement with a PMI. Finally, the efficiency of a MI extends to very high frequencies, finally limited by the absorption of the beamsplitter. Therefore for many measurements, low pass filters are required in order to prevent aliasing. The efficiency of a PMI is a natural low pass filter with a cutoff at ν_{max}. For some measurements, no additional low pass filter will be needed.

Another difference between the two interferometers is the offset term. If we consider a frequency distribution $I_o(\nu)$ rather than a monochromatic frequency, the outputs of the instruments are given by integrals over frequency.

$$I_{MI}(\Delta) = \frac{1}{2} \int I_o(\nu)\varepsilon_{MI}(\nu)[1 + \cos\Delta]d\nu \quad , \tag{37a}$$

$$I_{PMI}(\Delta) = \frac{1}{2} \int I_o(\nu)[\varepsilon_{PMI}(\nu) \cos\Delta + a(\nu)]d\nu \quad . \tag{37b}$$

In the limit $\Delta\to\infty$, the terms involving $\cos\Delta$ vanish and the residual intensity $I(\infty)$ is due only to the offset term. In a MI the offset term is equal in magnitude to the $\cos\Delta$ term at $\Delta=0$. This has two consequences. The available dynamic range of the data acquisition system is reduced. A much more important consequence, however, is the effect of source fluctuations. Suppose that we write the output of

0 0 0 0 4 3 0 8 0 1 2

the light source as $I_o(1 + \eta f(t))$, where $f(t)$ represents the source fluctuations. If the average mirror velocity is v, then in the measurement of the interferogram these time-domain fluctuations appear as path-dependent signal. Using (37a) the interferogram in the presence of fluctuations is

$$I(\delta) = I_{MI}(\delta) + \frac{\eta}{2} f(\frac{2}{v}\delta) \int I_o(\nu)\varepsilon(\nu)d\nu$$
$$+ \frac{\eta}{2} \int I_o(\nu)\varepsilon(\nu)\ f(\frac{2}{v}\delta)\ \cos 2\pi\nu\delta d\nu\ , \tag{38}$$

where I_{MI} is the interferogram without fluctuations. If the Fourier transforms of $f(t)$ is $F(\omega)$, then the spectrum $B'(\nu)$ measured in the presence of fluctuations is

$$B'(\nu) = I_o(\nu)\varepsilon(\nu) + \frac{\eta}{2} F(\frac{v}{2}\nu) \int I_o(\nu)\varepsilon(\nu)d\nu$$
$$+ \eta F(\frac{v}{2}\nu) * I_o(\nu)\varepsilon(\nu)\ , \tag{39}$$

where the symbol * means convolution. As we see, source fluctuations enter in two different ways. The term $\frac{\eta}{2} F(\frac{v}{2}\nu) \int I_o(\nu)\varepsilon(\nu)d\nu$ is due to the offset. The term $\eta F(\frac{v}{2}\nu) * I_o(\nu)\varepsilon(\nu)$ is due to fluctuation effects in the amplitude of path-dependent interferogram structure. This term normally provides a smaller contribution to $B'(\nu)$ than the offset term. We may see this using the following argument. Suppose that $f(t)$ is slowly varying. Then $F(\frac{v}{2}\nu)$ will have sharp structure. However, the spectrum $B(\nu) = I_o(\nu)\varepsilon(\nu)$ will often be a slowly varying function of ν. The convolution will have the effect of integrating

several cycles of F with the smooth function B. As a result, much of the structure of F will be cancelled out in the convolution, and this term will often be a small contribution to $B'(\nu)$. We conclude that the offset term is the more important source of errors in the interferogram. While fluctuation effects are important in the laboratory, they are crucial in astrophysical measurements. Since the offset of a PMI is much less than a MI, fluctuation effects are substantially reduced.

As a final point of comparison between interferometers we consider double beam use. As was shown earlier the interferogram from a double-beam MI with a sample of transmission T in one beam is

$$I_{MI} = I_o \{1 - \frac{1}{2} \varepsilon_{MI}(1-T) - \frac{1}{2} \varepsilon_{MI}(1-T) \cos\Delta\} \quad . \tag{40}$$

In a PMI the beams which are reflected and transmitted by the output polarizer are collected at the detector. Suppose that the sample of transmittivity T is placed in the transmitted beam. Using Eq. (23), the component at 2ω is

$$I_{PMI} = \frac{I_o}{2} (1-T) \{\varepsilon_{PMI} \cos\Delta + a\} \tag{41}$$

It is apparent from examination of Eqs. (40) and (41) that all of the earlier remarks concerning the relative advantages of a PMI compared to a MI apply. The offset term, however, deserves special comment. It T is nearly unity, $1-T$ will be small. The MI output will consist of a large offset with a small path-dependent modulation. The PMI output, however, will always have an offset which is small compared

0 0 0 0 4 3 0 8 0 1 3

to the path dependent term. It seems clear that the PMI provides a far superior double-beam capability from the standpoint of both source fluctuations and data acquisition.

E. Alignment

In a conventional MI, alignment is a fairly simple procedure. The beamsplitter and mirrors have rotational symmetry about the normals to their surfaces, so that there are only two degrees of freedom which need to be adjusted to align the instrument. In a practical instrument these degrees of freedom are usually arranged to be two orthogonal tilt adjustments of the fixed mirror.

In a PMI both the angular properties of the beams and the their polarization states need to be adjusted in order to align the instrument. The orientation of P_1 with respect to the beamsplitter affects the polarization states of the instrument and must be adjusted. Also as we shall see below, the use of dihedral rather then flat mirrors increases by one the number of adjustments which must be made of the fixed mirror. Thus there are four adjustments in all to be considered.

We treat the effect of a misadjustment of P_1 in the limit that the polarizers and beamsplitter are perfect. The input polarizer P_1 is misadjusted by an angle ϕ so that P_1 transmits $\frac{1}{\sqrt{2}}[\hat{i}(\cos\phi-\sin\phi) + \hat{c}(\cos\phi+\sin\phi)]$. Also we allow the phase of the output signal at 2ω to be changed by an angle ψ. The analysis is straightforward and the output signal at 2ω is $\frac{I_o}{2}(\cos2\phi\cos2\psi\cos\Delta-\sin2\phi\sin2\psi)$. Adjusting the phase so that

$$\psi = -\phi \quad , \tag{42}$$

we find that the output of the instrument is

$$I = \frac{I_o}{2} \left(\cos^2 2\phi \cos\Delta + \sin^2 2\phi\right) . \tag{43}$$

For a small misalignment ϕ, the interferogram is degraded by

$$\eta_P \approx 1 - 4\phi^2 . \tag{44}$$

For $\eta_P = .99$, the polarizer need only be aligned within $\pm 3°$, which is not a difficult tolerance to achieve.

The dihedral or corner mirror is made of two flat mirrors which meet along a line called the axis. The angle between the mirrors should be 90°. A deviation from that angle (Fig. 8(a)) will be called an angle deviation θ. The axis should lie in the plane normal to the direction of propagation of the incident light. It the axis is out of this plane, the misalignment will be called a pitch deviation denoted by ϕ_P (Fig. 8(b)). The axis should be inclined at an angle of 45° with respect to the polarization vector of the incident radiation if it is to correctly rotate the plane of polarization by 90° upon reflection. Misinclination of the axis will be called roll deviation, denoted by ϕ_r (Fig. 8(c)). Finally the vector which bisects the dihedral corner angle and is perpendicular to the axis should be parallel to the incident radiation direction. Misalignment of this vector, which is equivalent to rotation of the mirror about the axis, is called yaw deviation (Fig. 8(d)) and is denoted by ϕ_y. We shall see that instrumental properties are unaffected by yaw.

0 0 0 0 4 3 0 8 0 1 4

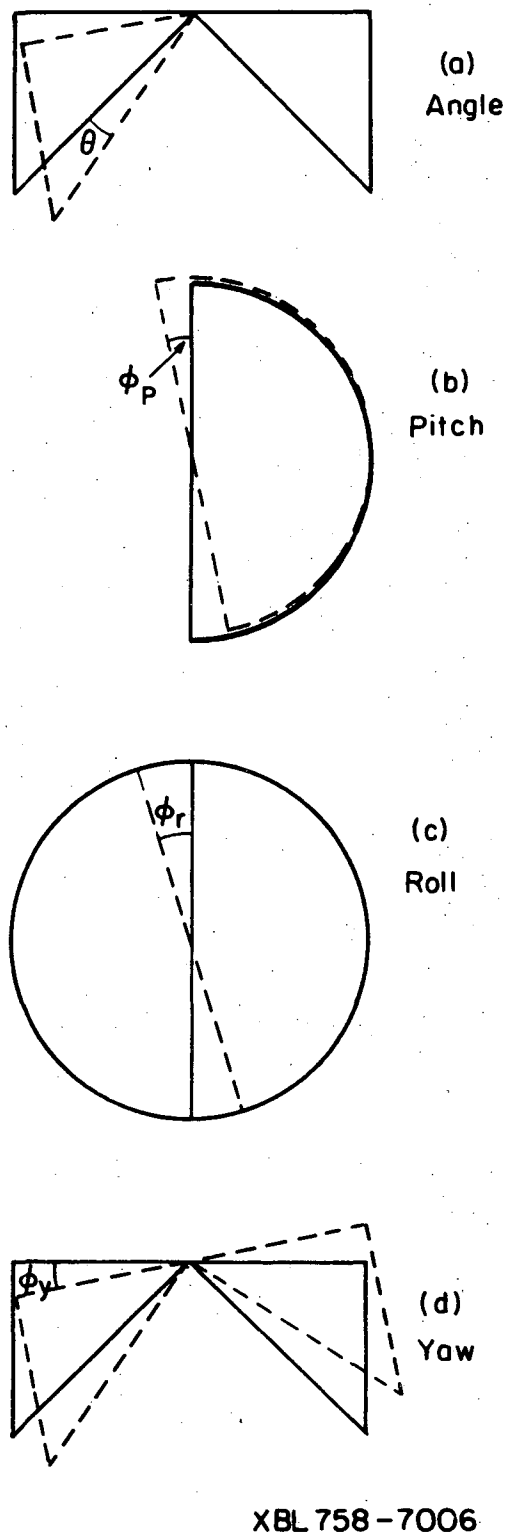


Fig. 8. Dihedral Mirror Misalignment.

We will first consider a correctly oriented dihedral mirror and assume that it is a perfect conductor. The boundary conditions for the reflection of radiation from a perfectly conducting surface are that the component of the electric field parallel to the surface and the component of the magnetic field perpendicular to the surface vanish. The incident radiation, polarized at an angle of 45° with respect to the dihedral axis, may be split into two equal components. One component (the TE mode) has the electric field oriented parallel to the axis while the other has the magnetic field oriented parallel to the axis (the TM mode). The electric field in the TE mode suffers a phase shift of π at the first surface and a second phase shift of π at the second surface because of the boundary condition on $E_{\parallel}$. Hence, E is unchanged upon reflection from the mirror. The magnetic field in the TM mode is not phase-shifted by the two reflections. The electric field is shifted by π. This can most easily be seen by noting that the propagation vector k, E, and H form a righthanded coordinate system. Since k is reversed by the corner mirror and H is unchanged, E must also be reversed in direction. Adding the reflected TE and TM modes, the total E vector is now perpendicular to the incident E vector. The plane of polarization has been rotated by 90°.

We will now consider in order of increasing complexity the effect of the four independent perturbations of the dihedral mirror on the performance of a PMI. The simplest perturbation is yaw. We first observe that since the dihedral angle is 90°, the reflected beam will be parallel to the incident beam. Next, by the use of a geometric

0 0 0 0 4 3 0 8 0 1 5

construction and some complicated but straightforward manipulation one can show that yaw does not affect the length of the path traversed by the incident beam. Finally the presence of yaw does not change the polarization argument presented above in any way. The only effect that yaw has on PMI performance is a minor geometric loss. If the dihedral mirror, which we assume to be circular with radius R is filled by the beam incident upon it, then yaw will lead to a reduction of the effective area of the mirror. The fraction lost is just $\cos\phi_y$. The efficiency of the PMI is then $\cos\phi_y$. No offset term is introduced by the beam imbalance. This weak perturbation of ideal PMI response can, of course, be eliminated by use of oversized mirrors.

The next perturbation we will treat is roll. It is clear that since the geometric properties of the beam striking the mirror are axially symmetric about the roll axis, roll cannot affect the geometric properties of the beam. Hence the only effect of roll is upon the polarization state of the reflected radiation. If the roll angle is ϕ_r, then relative to the mirror axis the incident radiation is polarized at an angle of $45° - \phi_r$. Resolving the radiation into TE and TM components as before, we find that the TE component has amplitude $\frac{E_o}{\sqrt{2}}(\cos\phi_r + \sin\phi_r)$ and the TM component has amplitude $\frac{E_o}{\sqrt{2}}(\cos\phi_r - \sin\phi_r)$, where E_o is the amplitude of the incident wave. Upon reflection, the TM wave has been phase-shifted by π and is therefore $\frac{E_o}{\sqrt{2}}(\sin\phi_r - \cos\phi_r)$. Recombining the reflected components, we find that the plane of polarization has been rotated by $90°-2\phi_r$. Using this result to solve for the output of an otherwise perfect PMI

with roll, we find that the signal at 2ω is

$$I = \frac{I_o}{2} (\cos 2\phi_r \cos\Delta + \frac{1}{2} \sin 4\phi_r) \ . \tag{45}$$

An efficiency of .99 can be achieved if $\phi_r < 4°$. The offset at that roll angle will be about 14% of the modulation term. It is clear that the proper adjustment of roll is more important for reducing the offset signal than for improving the PMI efficiency.

The treatment of a pitch misalignment is more difficult. If the pitch angle is ϕ_p, then the reflected beam is inclined at $2\phi_p$ with respect to the incident beam. This gives rise to two effects. First the reflected beam is no longer in the plane of the interferometer and there will be vignetting losses. A much more important effect, however, is a difference in path length across the mirror. There is also a minor polarization effect due to consideration of the component of the polarization of the reflected beam in the direction of the incident beam. We will calculate a conservative estimate of the effect of misalignment by considering only the variation of path length in the limit of small ϕ_p. We shall follow the calculation of Williams[11] and consider a circular mirror of radius R. We note that if we neglect polarization effects (in the spirit of this calculation), a pitched dihedral mirror and a tilted flat mirror are equivalent. If the mirror were perfectly aligned, then the contribution to the output signal from an area dA would be $\frac{I}{2A} \cos\Delta dA$, where as usual, $\Delta = 2\pi\nu\delta$. We consider an element of area $dA = 2(R^2-x^2)^{1/2}\ dx$. The path length for this element of area is $\delta + 2\phi_p x$. The total signal is, then,

0 0 0 4 3 0 8 0 1 6

$$I(\delta) = \frac{I}{\pi R^2} \int_{-R}^{R} (R^2-x^2)^{1/2} \cos[2\pi\nu(\delta+2\phi_p x)]dx \ . \tag{46}$$

Integrating[12] we get

$$I(\delta) = \frac{I}{2} \frac{2J_1(a)}{a} \cos\Delta \quad , \tag{47}$$

where J_1 is the usual Bessel function and

$$a = 4\pi\nu\phi_p R \quad . \tag{48}$$

Writing $\varepsilon_p = \frac{2J_1(a)}{a}$ and expanding the Bessel function for small a,

$$\varepsilon_p \approx 1 - \frac{a^2}{8} \quad . \tag{49}$$

For two-inch optics, the efficiency ε_p will be .99 at 100 cm^{-1} for $\phi_p \approx .13$ milliradians. This tolerance is exceedingly stringent and great care must be taken if the alignment is to be maintained to this precision. The tilt of the mirrors of a MI must also be adjusted to this precision.

The final and most complicated aberation is angle aberation. When the dihedral angle is not 90°, incident rays are reflected at an angle with respect to the input direction. Therefore we expect vignetting and path-length changes, as in the above case. There is an additional complication, however. Suppose that we consider the incident beam (assumed of uniform radiation density and circular cross-section) to be divided into two half-beams of semicircular cross-section, each of which is incident on one surface of the dihedral mirror.

Because of angle aberation, not only are the two reflected half-beams skew to the plane of the interferometer, but they are not parallel to each other. The two half-beams travel different length paths. We have calculated the path lengths for the two half-beams exactly. Qualitatively, the rays which strike the mirror at the edge travel different distances than the rays which strike near the axis. To order θ the change of path length is $2\theta x$ for both half-beams. The paths traveled by the two half-beams differ only to order θ^2. We can calculate the effect of these changes in path length by the same technique that was used above. We calculate the effect for the two half-beams separately and add. This accounts for the fact that the output angles of the two half-beams are different. The integral is

$$I(\delta) = \frac{I}{\pi R^2} \int_0^R dx \int_{-\sqrt{R^2-y^2}}^{\sqrt{R^2-y^2}} dy \cos(\Delta+4\pi\nu\theta y) \ . \tag{50}$$

Integrating, we find that

$$I(\delta) = \frac{I}{2}\left\{\frac{2J_1(b)}{b}\cos\Delta - \frac{4}{\pi}\sin\Delta \sum_{k=0}^{\infty} \frac{(-1)^k\, b^{2k+1}}{(2k+1)!!(2k+3)!!}\right\} \ , \tag{51}$$

where $b = 4\pi\nu\theta R$. The second term, which did not arise in Eq. (47), is due to a lack of cancellation across the mirror in this case. The small b approximation is

$$I(\delta) = \frac{I}{2}\left\{\left(1 - \frac{b^2}{8}\right)\cos\Delta - \frac{4b}{3\pi}\sin\Delta\right\} \ . \tag{52}$$

0 0 0 4 3 0 8 0 1 7

The second term introduces a phase error. To order b^2, we may rewrite (52) as

$$I(\delta) = \frac{I}{2}\left(1 - \frac{b^2}{8}\right) \cos(\Delta+\psi) \quad , \tag{53}$$

where

$$\psi = \frac{16}{3} \nu\theta R \quad . \tag{54}$$

Since ψ is frequency dependent, the phase error introduced is linear and therefore represents a shift in the zero-path point by $\frac{8}{3\pi}\theta R$. From Eq. (53) we see that the alignment criterion for angle aberation is as strict as for pitch aberation. If θ has been adjusted to .2 milliradians, the zero-path shift is 4 microns. We also note that both dihedral mirrors must be independently adjusted for angle misalignment. Unlike the other adjustments, an error in the angle of the movable mirror cannot be cancelled by an error in the angle of the fixed mirror.

F. Construction Details

The PMI which I constructed was based on a modified "cube" interferometer obtained from H.A. Gebbie. It consists of a cast aluminum cube in which the beamsplitter is mounted, with four arms attached to the cube faces. The four arms house the input optics, the fixed and movable mirrors, and the output optics. The cube and arms are sealed to form a single vacuum chamber. The interferometer is evacuated during use to eliminate the effects of atmospheric water vapor absorption. The interferometer is schematically indicated in Fig. 9.

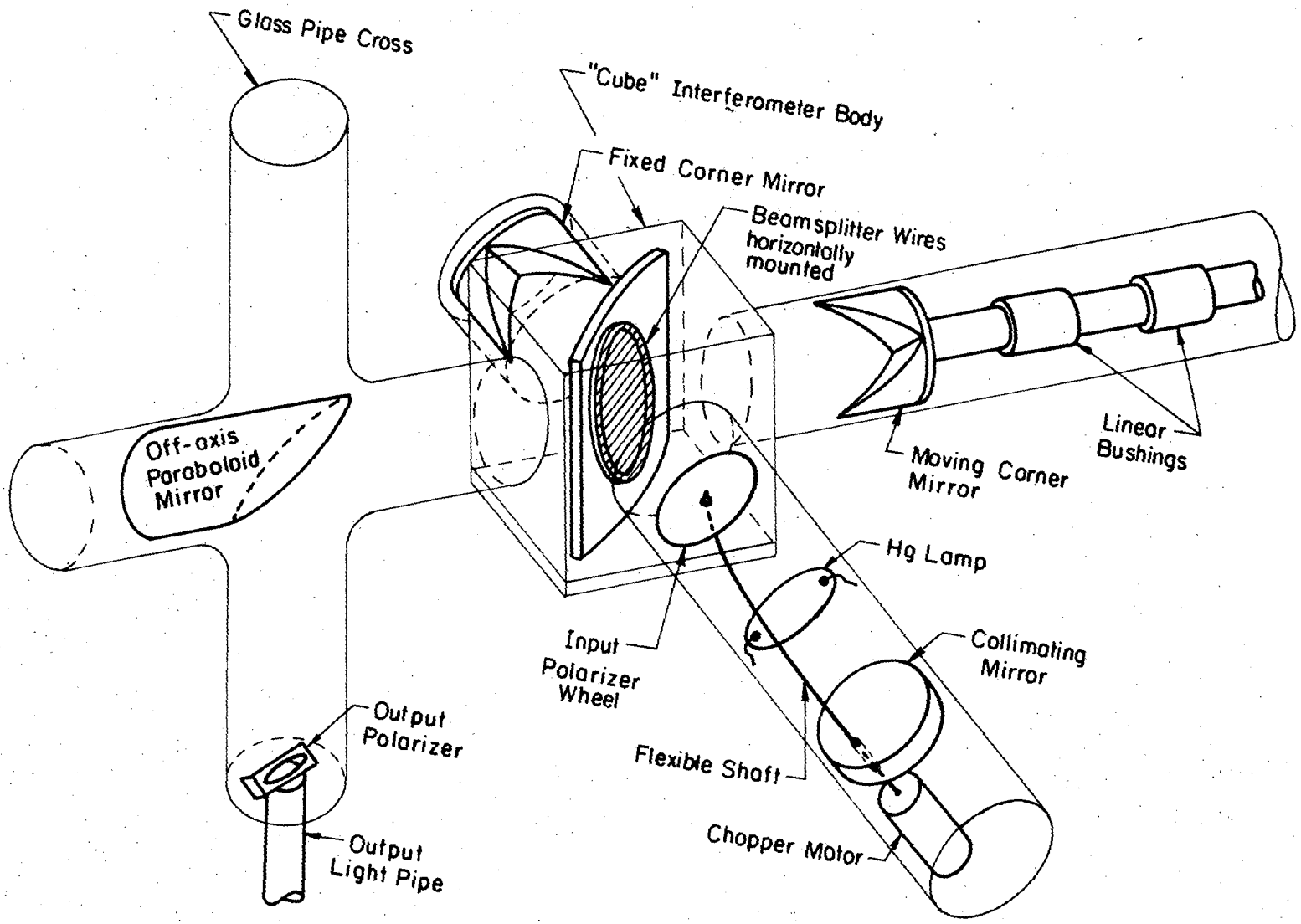


Fig. 9. Polarizing Michelson Interferometer.

XBL758-7007

0 0 0 0 4 3 0 8 0 1 8

The light source for the interferometer is the quartz discharge tube from a 100 Watt G.E. H-100 mercury arc lamp. The output of the lamp is collimated by an f/1.5 paraboloidal mirror. In order to eliminate ultraviolet radiation from the lamp, which is harmful to the Mylar substrate of the input polarizer, a stainless steel baffle was placed in front of the lamp to block direct radiation. In addition, the surface of the collimating mirror was lightly coated with flat black paint. The coating effectively eliminated reflection of ultraviolet radiation without appreciably affecting the far infrared output of the source optics. Excess heat was removed by using double-wall construction for the housing of the input-optics arm. A continuous flow of water was maintained between the walls. Heat transfer to the housing was provided by heavily coating the inside of the housing with black paint.

The input polarizer was also mounted in the input-optics arm. It consisted of a large light-weight phosphor-bronze wheel to which a polarizer was attached with epoxy. The polarizer, consisting of 1000 gold lines per inch deposited on a Mylar substrate, was made by the Buckbee-Meers Corp. For this instrument it was decided to rotate the input polarizer. A Globe 2-phase synchronous motor was mounted behind the collimating mirror. A flexible shaft made from stainless steel hypodermic syringe tubing came through a small hole in the mirror and connected the polarizer wheel to the motor. The bearing at the polarizer end of the shaft was made from Teflon and was designed to incline the polarizer by about 15°. This design was convenient

and compact in terms of construction, but in practice turned out to be a major weakness of the design. The shaft wobbled as it rotated causing polarizer vibrations. Also the flexing of the shaft caused the Teflon bearing to be loaded unevenly, leading to a very appreciable problem of wear.

The beamsplitter was mounted in the interferometer "cube". It consisted of a wire grid of .001 inch tungsten wires with a wire-to-wire spacing of .003 inches. As suggested by Vickers, et al.,[13] the grid was fabricated on a conventional lathe. A winding form with flat sides was chucked in the lathe, which was geared to rotate as slowly as possible. Wire was fed to the form from a tension-maintaining apparatus attached to the compound rest. Wire spacing was controlled by the longitudinal feed of the lathe. Wire was held in place during the winding procedure by G. E. varnish. This winding technique worked fairly well, but may not be suitable for winding thinner wire at finer spacing. The problem occurs at the edges of the winding form, where the tensions on the wire are largest.

Since when winding a beamsplitter, wires break most often at the edges of the coil form, an improved coil winder for high frequency beamsplitters should have more gentle bends at the edges. Suppose we wish to wind a beamsplitter for an interferometer with 3-inch optics. The beamsplitter would be about 5 inches in diameter. An improved coil form would consist of a cylinder 12 inches in diameter. One side of the cylinder would be milled flat over an area at least 5×5 inches. In cross-section the form would look like a circle with a 5 inch chord

0 0 0 4 3 0 8 0 1 9

removed. Since the bend at the transition from circle to chord is fairly gentle, it should be possible to wind high frequency beamsplitters using this form.

The beamsplitter was epoxied to a mounting ring which was then attached to the beamsplitter mount as shown in Fig. 10. This mounting procedure stretches the wires over the rounded edge of the mount. This edge was ground flat, to ensure the flatness of the stretched beamsplitter. Since the tensions involved in the beamsplitter are very large (roughly 1000 wires with a few ounces of tension on each wire), the mount and mounting ring were made of the thickest stainless steel that space considerations allowed. This was necessary to prevent beamsplitter warpage due to wire tension. The beamsplitter was mounted in the "cube" with a kinematic suspension so that no warpage could occur due to mounting of the beamsplitter.

The problem of component flatness, particularly beamsplitter flatness, is common to all interferometers. It is clear that if a component is not flat, the path lengths for beams which strike the component in different places will be different. The effect of component warpage will be qualitatively similar to the effect of dihedral mirror pitch described above. If the warpage is to have negligible effect on the efficiency of the instrument, then for 2-inch optics at 100 cm^{-1}, all components must be flat within .1 milliradians. Defining the warpage in terms of a radius of curvature, then we find that the radius of curvature must be greater than 10^4 inches.

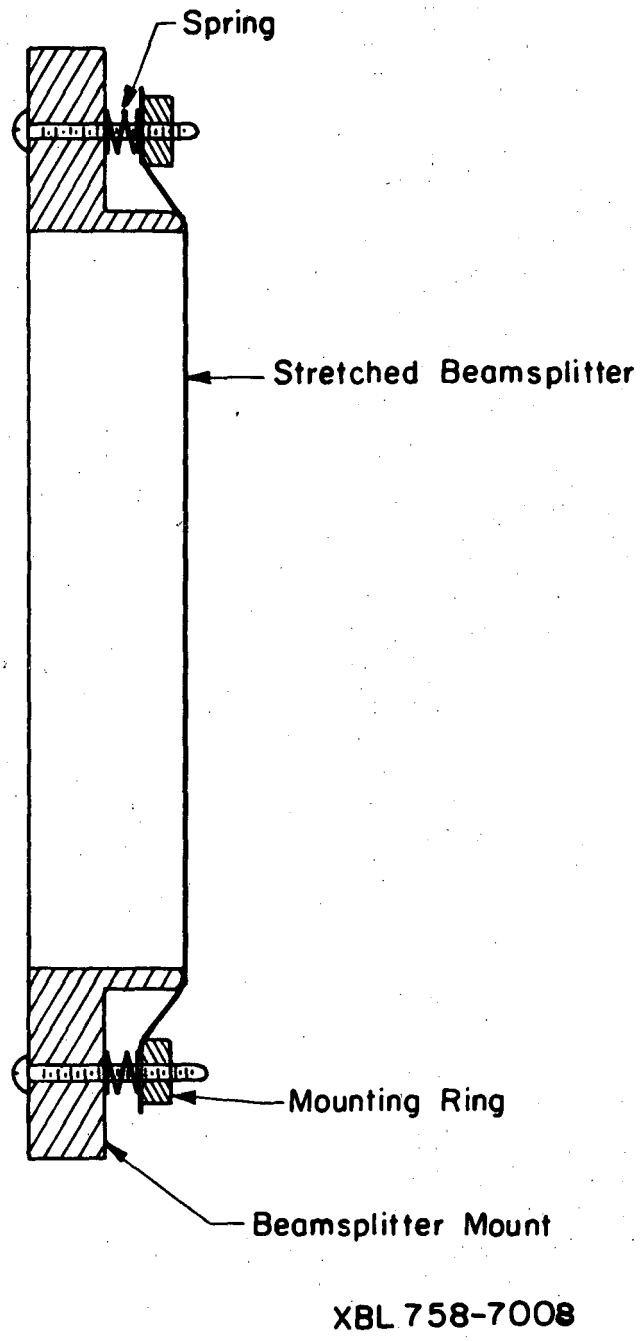


Fig. 10. Beamsplitter Mounting.

0 0 0 0 4 3 0 8 0 2 0

The moving corner mirror was mounted on the end of a push rod. The rod passed through two well-separated Thomson bushings which provided suspension for the movable mirror assembly. A micrometer pushed on the rod to provide the mirror motion. It was driven by a Slo-Syn stepping motor. The rod was held against the micrometer by a spring. The spring also provided a rotational torque which held the mirror against a guide rod so that the mirror would not roll as it moved.

The fixed mirror was mounted using a modification of the standard three-point mirror mount. The mount was designed so that the angle and pitch aberations of the mirror could be controlled by small micrometers which extended through the vacuum housing. Therefore these adjustments could be made while the interferometer was evacuated. Roll aberation was corrected by an overall rotation of the fixed mirror housing. Since the positive drive direction of the movable mirror is toward the "cube", and since it is desirable to acquire data in the direction of increasing path length, the fixed mirror housing set the fixed mirror back sufficiently far that zero path difference occurred near the fully-retracted end of the movable-mirror travel.

The output optics arm contains the output polarizer and the optics required to focus the 2-inch parallel beam from the interferometer into standard $\frac{1}{2}$ - inch light pipe. The focusing was accomplished using an f/1.5 off-axis paraboloidal mirror. This mirror focuses parallel light incident along the mirror axis to a spot along a perpendicular to the mirror axis, thus changing the direction of

propagation of the radiation by 90°. This type of mirror was used to simplify the requirement that the output of the PMI be physically consistent with the output of the MI normally in use. The mirror was made by mounting a cylinder on the faceplate of a lathe a focal length off-axis. The surface of this cylinder was then machined to computer-generated surface coordinates. After polishing, the mirror was found to focus visible light to a 2mm diameter focal spot.

The PMI was aligned using procedures developed by John Mather[14] and tested. The instrument displayed a cutoff frequency of about 80 cm^{-1}, consistent with the grid spacing limit. Vignetting was shown not to be serious by measuring the spectrum of water vapor. Since water vapor lines are much narrower than the .1 cm^{-1} resolution limit imposed on the PMI by the length of the micrometer screw, one expects significant modulation of the interferogram of water vapor to extend to maximum path length. This was shown to be the case. The PMI was used extensively in the measurements described in Chapter IV.

Although it has never been tested, a double-beam modification of the PMI has been constructed. The output polarizer is replaced by the apparatus shown in Fig. 12. The output polarizer is inclined at 45° so that the transmitted and reflected beams are perpendicular. Both beams are reflected once and converge on the beam-recombiner. The sum of the two beams emerges from the bottom of the device. Sample holders may be mounted in both beams so that the difference in the transmission of the samples can be measured. Alternately two reflecting samples may replace the mirrors and the difference in reflectivity may be measured.

0 0 0 0 4 3 0 8 0 2 1

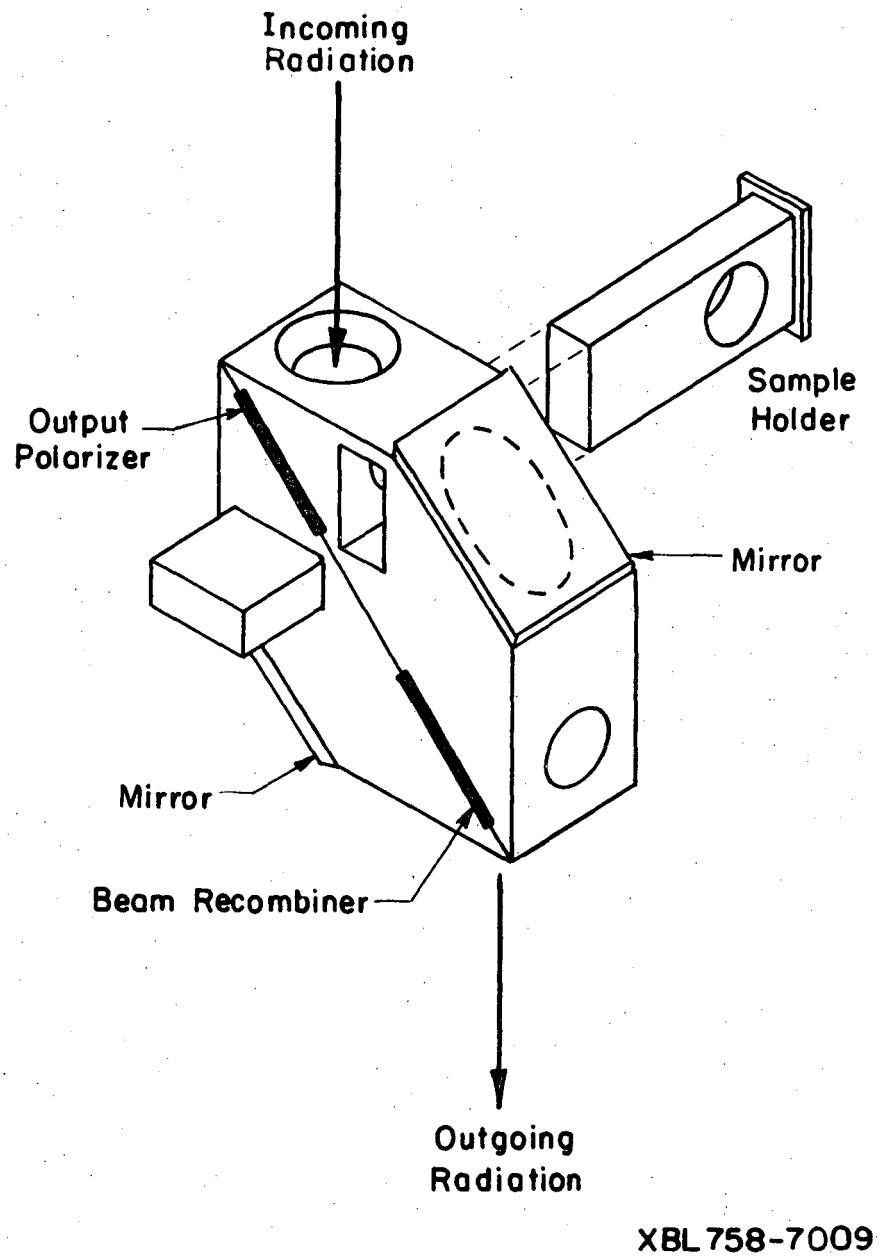


Fig. 11. Double Beam Apparatus.

CHAPTER II REFERENCES

1. D. H. Martin and E. Puplett, Infrared Phys. 10, 57 (1966).
2. D. P. Woody, J. C. Mather, N. S. Nishioka, and P. L. Richards, Phys. Rev. Letters 34, 1036 (1975).
3. E. V. Loewenstein and A. Engelswrath, Air Force Cambridge Research Laboratories Technical Report AFCRL-66-536 (1966).
4. M. Born and E. Wolf, Principles of Optics, (Pergamon Press, Oxford, 1965), p. 323 ff.
5. Ibid., p.36 ff.
6. N. Marcuvitz, ed., Waveguide Handbook, (McGraw-Hill Book Co., New York, 1951), p.280 ff.
7. K. D. Möller and W. G. Rothschild, Far-Infrared Spectroscopy, (Wiley-Interscience, New York, 1971), p. 92 ff.
8. N. Marcuvitz, op.cit., p. 285 ff.
9. N. Marcuvitz, op.cit., p. 285 and pp. 268-271.
10. J. P. Casey and E. A. Lewis, J. Opt. Soc. Am. 42, 971 (1952).
11. C. S. Williams, Appl. Opt. 5, 1084 (1966).
12. Use was made of Gradshteyn and Ryzhik, Table of Integrals, Series, and Products, (Academic Press, New York, 1965).
13. D. G. Vickers, E. I. Robson, and J. E. Beckmen, Appl. Opt. 10, 682 (1971).
14. J. C. Mather (Ph.D. Thesis), Lawrence Berkeley Laboratory Technical Report LBL-2258 (1974).

0 0 0 0 4 3 0 8 0 2 2

III. INSTRUMENTATION FOR FOURIER SPECTROSCOPY

In an experimental study of some sample property (for example, sample transmission), the interferometer (such as the PMI described in the previous chapter) is only the first step in the measurement process. The output of the interferometer is applied to the sample, and the transmitted radiation is synchronously detected. The output of the lock-in amplifier is a high level (typically 10 Volts full scale) signal proportional to the power falling on the detector. Before the advent of minicomputers, the data were typically digitized and recorded on punched cards, magnetic tape, or some similar medium. Fourier transformation and data analysis were performed off-line at a centralized computer facility.

Small, low-priced, but powerful minicomputers have changed the way that Fourier Transform Spectroscopy (FTS) is done. Data from the lock-in amplifier are digitized and fed directly to the minicomputer. The data are stored in computer memory and are instantly available for real-time data analysis. The computer also controls various aspects of the experiment such as mirror motion. In this chapter I will describe the development of a computerized instrumentation system primarily intended for FTS, but with much more general capabilities. The first section will discuss what I call the data acquisition system (DAS). This is the part of the instrumentation system which connects directly with the experiment. Next I will discuss the general problem of computer interfacing and describe how the DAS and other system peripherals are connected to the computer. The third section will

deal with the software operating system written for FTS.

A. Data Acquisition System

The data acquisition system (DAS) consists of five nearly independent units, each of which performs one aspect of the data acquisition process. These units are the analog section, the external interrupt section, the clock, the stepping motor control, and the display unit. The five units are housed in a single enclosure which requires 8 $\frac{3}{4}$ inches of standard 19 inch relay rack space. The enclosure also contains a power supply which provides +5 VDC at 6 amps and ±15 VDC at 1 amp each. Surplus power (5V at 2.5 amps and ±15 V at .5 amps) is available on the rear panel. Air flow for cooling the electronics is also provided. Air enters the unit and passes around and through the analog section, which is the unit most sensitive to temperature variations, moves through the box and is pulled out of the box by a fan. As the air leaves the box, it passes last over the power supply, which is the major source of heat in the system. This pattern of air flow helps isolate power supply heat from the rest of the system. In addition to the surplus power connector, the cables which connect the DAS to its computer interface (three 40-wire flat cables) and the cable which connects the stepping motor control to the Slo-Syn Preset Indexer (Superior Electric Co.) which runs the stepping motor also attach to the rear panel. All other connections are made to the front panel.

0 0 0 4 3 0 8 0 2 3

The DAS was designed with flexibility of operation in mind. Although the system is primarily intended for use in FTS, other experiments may be of interest from time to time. This flexibility to perform various kinds of measurements is provided by allowing the computer to control the most important features of the data acquisition process. By the use of suitable software, for example, the DAS might easily be made to operate as a multi-channel analyzer, a boxcar integrator, a transient digitizer, or some similar instrument. It will not be capable of the high speed of operation achievable with special purpose instruments because of the necessary computer intervention in the acquisition process, but for measurements requiring acquisition rates up to 60 kHz or so, it will be able to replace any of these special purpose instruments.

A block diagram of the analog section is shown in Fig. 12(a). The analog signal to be measured is selected by the analog multiplexer, amplified, and sampled by the sample-and-hold amplifier. The sampled signal is digitized by the analog-to-digital converter (ADC) and the digital output is transmitted to the computer.

The input signals enter the box via insulated BNC connectors. The shells are connected to ground through 10Ω resistors to help reduce possible ground loops. The analog section is protected against overvoltages by the diode clamp circuit shown in Fig. 12(b). The effective leakage resistance of these diodes is about 10^{12} Ω, and simple equivalent circuit calculations show that for signals of magnitude less than 10V there is no measureable non-linearity introduced.

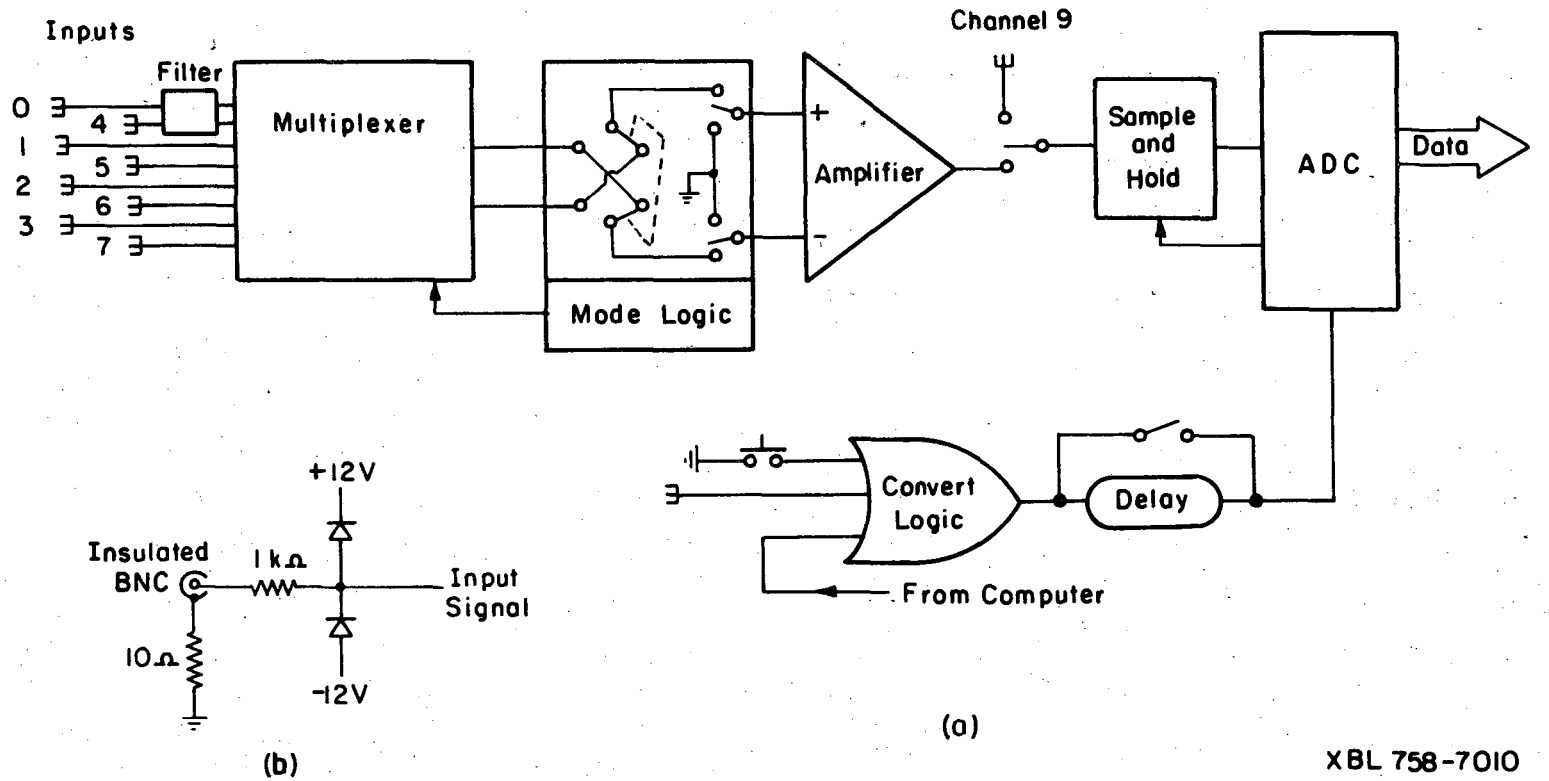


Fig. 12. DAS Analog Section--Block Diagram.

0 0 0 4 3 0 8 0 2 4

The 1 kΩ 1/8 Watt series resistor acts like a fuse to protect the diodes from drawing too much current. The circuit will protect the analog section continuously against voltages up to 25V. Beyond this point, the series resistor will fuse, but the inputs will continue to be protected. Input voltages over 200V will damage the diodes.

The multiplexer will handle either eight single-ended inputs or four differential inputs, depending on the position of a panel switch. In differential mode, channels 0 and 4, 1 and 5, etc. are the input pairs. Channels 0 and 4 incorporate low pass input filters with switch-selectable cutoffs. The multiplexer channel is selected by a ten-position thumbwheel switch. Positions 0 through 7 select the respective input channels. Position 8 allows the computer to select the channel. Position 9 selects a direct input channel that bypasses the multiplexer and amplifier. In addition to the choice of single-ended or differential operation, there exists mode-control logic, which may be used to invert the input signal, ground it, or when in differential mode ground one or the other side of the differential output. These modes are controlled at the front panel.

The amplifier is a gain-switchable instrumentation amplifier designed for operation to over 100 kHz. Gains of 1, 2, or 10 may be selected, which allows full scale digital outputs of the ADC to correspond to 10V, 5V or 1V respectively. The gain is controlled by a front panel switch. An offset adjustment is also provided.

The sample and hold amplifier (SHA) is a fast gain 1.0 amplifier with a storage capacitor on the output. In sample mode, the output voltage follows the input voltage. In hold mode the amplifier output is disconnected from the capacitor and the output of the SHA remains at the voltage level reached at the time the control signal changed to hold-mode. In use the slow discharge of the capacitor due to leakage resistances is negligible. The purpose of this device is to provide a constant voltage to the ADC during the ADC conversion time even in the presence of time-varying inputs. A change in the ADC input level during a conversion will lead to an erroneous digital output.

The ADC performs an analog-to-digital conversion with a resolution of one part in 2^{15} in 10 μsec. This device (model ADC 1251 manufactured by Phoenix Data Corp.) achieves this performance by means of a novel variation of conventional successive approximation techniques. The result is a 16-bit digital result guaranteed accurate to better than 15 bits. This conversion technique offers two advantages at the price of increased complexity. This ADC is able to detect the presence of a voltage greater than the full scale value. This indicator, which is not normally provided by ADC's, is useful as an error indicator. Also the nature of the conversion process requires that the digital output be latched in an internal register. This means that the output of one conversion is available for use until the end of the next conversion. In most ADC's, the output is only available until the start of the next conversion. A simplification of the external circuitry required to use the device is achieved.

A conversion may be initiated in one of three ways. First the computer may issue a command-to-convert (CTC) signal. Second, a front panel push button may be the source of CTC. Finally the leading edge of a TTL-compatable external signal, input at the front panel, may cause a CTC to be issued. If desired, the CTC may be delayed by a variable time delay. This delay might be useful, for example, if the DAS were being used as a box-car integrator. In this mode of operation, the trigger pulse is applied to the external convert input and the delay set to select the point on the analog waveform to be measured. The integration of the signal would be done digitally in the computer.

In many types of experiment it is important to notify the computer that some event has taken place. The external interrupt section of the DAS provides this capability. Basically this section is a voltage discriminator triggerable on either the leading or the trailing edge of the signal. Suppose the edge-select has been set for leading edge triggering. Then whenever the input voltage changes from less than the trigger level to greater than the trigger level, a pulse is issued which causes a computer interrupt. Interrupts will be discussed below. The trigger level is adjustable from -15V to +15V. In order to help set the trigger level, the leading decimal point of the display (discussed below) is flashed whenever an interrupt is generated by the section. A front panel switch inhibits interrupt generation.

For some experiments, synchronization of the computer with some part of the experiment is useful. The clock section provides this capability. Frequency division of a 2 MHz crystal oscillator provides square wave outputs of 10 kHz, 3.333 kHz, 1 kHz, 333 Hz, 100 Hz, 33 Hz, and 10 Hz. In addition an external input is provided. The frequency desired is chosen by a thumbwheel switch. The selected frequency is applied to the external input of the KW11-P real-time clock. This device, described in Appendix A provides timing information and control to the computer. The selected frequency is also available as a front-panel output for use in synchronizing an experiment.

The stepping motor used to move the mirror of an interferometer is controlled by a Slo-Syn Preset Indexer. It has two important modes of operation. If commanded to perform a Jog, the Indexer advances the motor one step. If commanded to perform an Index, the Indexer advances the motor the number of steps set on dials on the Indexer. Motion either clockwise or counterclockwise is possible. Computer control of the Indexer is achieved by the stepping motor control section of the DAS, coupled with internal modification of the Indexer. Index or Jog pulses from the computer or from pushbuttons on the front panel cause appropriate pulses to be issued to the Indexer, initiating the operations. Direction is also controllable from either the computer or the front panel. A switch also allows these control pulses to be inhibited. This circuitry need not be restricted to stepping motor control, however. The rear panel output connector provides a ground and four signal lines. Two of these lines carry TTL-compatable Index

0 0 0 0 4 3 0 8 0 2 6

and Jog pulses. One carries the direction logic-level. The last carries the device response pulse. When either Index or Jog is done and the device is ready to accept another Command pulse, it issues a Ready pulse. Any device which can accept Command pulses and issue Ready pulses can be attached in place of the Indexer.

In order to facilitate operator use of the DAS, a numerical display is provided. It is capable of accepting 16-bit binary information (in 2's complement arithmetic) and displaying this information in the form of a signed decimal integer in the range from -32768 to +32767. The binary input information may either come from the ADC output or from the computer. The display is updated ten times per second so that time variations of the input data are readily apparent. In addition a numeral one located immediately to the right of the sign character is used to indicate ADC overrange and a decimal point is used to indicate that an external interrupt has occurred. The conversion from binary- to decimally-coded information is done in a simple and straightforward manner. Suppose the binary number is positive. It is loaded into a binary counter while a decimal counter is set to zero. The binary counter is clocked down to zero while the decimal counter is clocked up. Reaching zero halts the clock. The number residing in the decimal counter is easily displayed. Negative numbers are handled similarly except the binary counter is now clocked up until it overflows.

B. Peripherals and Interfacing

A computer, by itself, has no way of inputting or outputting information. The devices to which it is connected which allow it to communicate with the outside would are known as peripherals. In this section we will discuss the peripherals which are connected to our PDP 11/20 computer and the way in which that connection or interface occurs. In order to do this, we must first discuss a little about the architecture of the PDP 11 series of computers. For more detailed information the reader is referred to the numerous handbooks and manuals published by Digital Equipment Corp. (DEC).

In the PDP 11, all communication between the central processor and peripherals and between peripherals occurs over a single multi-conductor cable known as the Unibus.[1] The Unibus consists of three types of signal lines. First is the address lines. Since every word of memory and every aspect of peripheral operation is treated similarly by the Unibus, some way is required to distinguish between all of the different pieces of information available. This is done by giving everything (except for central processor internal states) addresses. Every word of memory has a unique address. In order to communicate with a peripheral, each peripheral is also given one or more unique addresses. These addresses are called peripheral registers. The address lines are used to select what information is to be carried on the Unibus. The data lines carry that information. The control lines handle timing and control the type of information transfer. Three kinds of information transfer will be important for

0 0 0 0 4 3 0 8 0 2 7

this discussion. The process by which the processor obtains information from a peripheral is a read operation. The process by which the processor outputs information to a peripheral is called a write operation.[2] If the processor requires a word of data stored in memory in order to perform a calculation, it reads that word of memory. After completing the calculation, if the result needs to be stored for later use, the processor writes it in an appropriate memory location. The ADC output of the DAS, as another example, is assigned a particular peripheral register. In order to obtain the result of a measurement the processor reads that register. If the information is to be stored for later use, it is then written into some word of memory.

The final kind of bus operation we will be concerned with is the interrupt. Since most input/output operations are slow compared to the speed of instruction execution in the processor, it is highly inefficient for the processor to continually read peripheral registers to see if some operation such as typing a character on the teletype has been completed. At the same time, however, the processor must be notified upon completion of such a task so that it may issue the command to perform the next task. This problem is resolved by allowing the processor to continue executing whatever calculations are called for and then requiring that the peripheral notify the computer when it is done with its task. The calculation is then interrupted and a special set of instructions are executed to service the peripheral. This sequence is known as an interrupt. When the peripheral finishes its task, it issues an interrupt request on the appropriate control line. The

computer architecture includes a priority structure so that some requests may take priority over others, depending on the importance and immediacy of the information. For example, the request associated with completing an ADC conversion has a higher priority than the request associated with typing a teletype character since the former is more important. When the processor grants the interrupt request, the peripheral places an address on the Unibus known as the vector address. This is an address in memory which contains the address of the first instruction of the code which services the interupt. The processor suspends normal operation, stores the information it needs to return to the calculation in progress at the time of the interrupt, and begins executing the first instruction of the service routine. When finished with the service routine, the processor resumes execution where it left off. Each different interrupt request must have a unique vector address.

From the above description of the system architecture, it is clear what must be done to interface a peripheral to the Unibus. The Unibus and the peripheral each have their own set of signals and protocols for information transmission. The interface must make them compatable. It must recognize the addresses assigned to the peripheral and connect the appropriate signals to the data lines at times governed by the Unibus control lines. It must also recognize ready signals from the peripheral and be able to use them to generate interrupt requests and the associated vector addresses.

0 0 0 0 4 3 0 8 0 2 8

Table 2 is a list of the peripherals presently interfaced to the PDP 11. From the viewpoint of software implementation, the interfacing of these peripherals requires knowing the addresses of the peripheral registers, what information is assigned to each bit of each register, what operations lead to interrupts, and what the priority and vector address of each interrupt is. All of this information is listed in Appendix A. Fortunately, all standard peripheral devices sold by DEC come equipped with interfacing circuitry. Only custom-designed peripherals require custom-designed interfaces, and DEC sells circuit boards which make interfacing many custom peripherals easy.

Four peripherals which I interfaced to the computer are the oscilloscope, the DAS, the card-punch, and the X-Y plotter. Of these, the scope was interfaced solely using standard circuits and will not be discussed here. A suggestion for an improved scope interface is the subject of Appendix B. The other three peripherals required custom design and will be discussed briefly.

The DAS interface was implemented using several standard DEC circuit cards and one custom-designed card. It provides three interrupts and contains four registers. The interrupt requests occur when an ADC conversion has been completed, when a stepping motor motion has been completed, and when an external interrupt occurs. Two of the registers are used for controlling the analog and stepper motor sections of the DAS and for monitoring the status of various operations. Control features include Index and Jog command bits, a stepping-motor direction bit, a CTC bit, and bits to enable the multiplexer channel to be

Table 2. Peripheral Devices

Peripheral	Description
Memory	12288 words of core memory
Memory	12288 words of semiconductor memory
KE11-A	Hardward Multiplication and Division (known also as EAE or Extended Arithmetic Element)
KW11-L	Line Frequency Clock
KW11-P	Programmable Real-Time Clock
KL11	Teletype
PC11	High Speed Paper Tape Reader and Punch
Oscilloscope	Displays data
DAS	Data Acquisition System
Card Punch	Punches alphanumeric information on computer cards
X-Y Plotter	Plots graphical information

0 0 0 0 4 3 0 8 0 2 9

selected by the processor. Status bits include flags which monitor stepping motor motion, ADC operation, and the overrange condition of the ADC. In addition there is a register for the ADC output and a register for display input.

The card punch interface consists of a status register, a data register, and one interrupt. The status register is used to monitor card punch operation and an interrupt is provided when the punching of a character is complete or when the card punch runs out of cards. The latter is an error condition. A character is punched by writing the standard ASCII code for the character into the data register. A specially programmed read-only memory is used to generate the Hollerith (card punch) code corresponding to that character. This code is used to drive the punch.

The x-y plotter interface allows the plotter to be run both directly and incrementally. In direct mode, the new coordinates of the pen are written into the x and y coordinate registers. This loads two 12-bit DACs with these numbers. The DACs produce voltage outputs proportional to these numbers, which are applied to the x and y inputs of the plotter, causing the pen to move. In the incremental mode, special bits in the control register may be set which will cause the numbers stored in the coordinate registers to be incremented or decremented by one. This will cause the pen to move by one small step up, down, left, right, or along the principal diagonals. Proper sequences of these incremental motions cause the pen to move in any chosen manner. An interrupt is available to indicate the completion of an incremental

motion. There is also a control bit to raise and lower the pen.

C. Computer Software - FTS-11

Having designed a data acquisition system and having interfaced it and other peripherals to a computer, all for the purpose of developing a computer-controlled FTS system, the next step is to write software to implement such a system. I will describe a special purpose assembly language program designed for Fourier transform spectroscopy called FTS-11. This program collects data, takes Fourier transforms, displays the data in various ways, outputs data for permanent off-line storage, and performs useful data analysis computations. Although the program is long and complicated it may be divided into six main sections for the purpose of discussion: the monitor, data acquisition, I/O, data analysis, buffer management, and utilities.

The heart of the system is the monitor. This section of the program is the overall scheduler and controller for everything the program does. It accepts operator commands input from the teletype, decodes them to determine what system response is called for, and transfers control to the piece of code required to execute the desired command. Whenever the monitor is in control of the system and is ready to accept a command, the program is said to be in command mode. The monitor indicates that command mode has been entered by typing an asterisk on the teletype. At this point the program will wait for the operator to enter the next command to be performed.

0 0 0 0 4 3 0 8 0 3 0

All commands are entered on the teletype. A command consists of one or two letters followed by up to four integers. The letters tell the program what to do. The numbers are parameters used to provide additional information. Some commands, which require more than four parameters, also ask for additional information interactively. Parameters if specified are separated from the letters and from each other by commas. A command is terminated by typing the RETURN key. This informs the monitor that the command is complete and may be executed. The monitor decodes the typed character string and transfers control to one of twenty-six pieces of code based on the first letter of the command. In general, the first letter determines the category of response called for, while the second letter specifies which of several possible responses in that category will actually be performed. Appendix C contains an abbreviated summary of the commands available to the operator.

In order to appreciate the way in which various operations are executed in this system, it is useful to define the concept of foreground and background tasks in a real-time environment. Data acquisition is an on-going real-time process. Thirty minutes or more may be required to acquire a single interferogram. Certain I/O operations may also take a substantial lengths of time. Yet these processes are characterized by the fact that relatively little actual computer time is required for their execution. Most of the time is spent waiting for things to happen. As was noted above, efficient use of the computer requires that these operations take place through the use of interrupts. We say

that when the computer is executing a peripheral service routine in response to an interrupt, it is executing a foreground task. All other tasks are called background tasks. Using this language we may describe, for example, the computer response to the command to acquire an interferogram. The monitor, in background, decodes the command and branches to the data acquisition routine. This routine sets up the necessary conditions for data acquisition, initiates the first ADC conversion, and returns to the monitor. The program, still in background, enters command mode and is available for the execution of other background tasks, such as analysis of data previously acquired and stored in the computer. In the meantime, the DAS will from time to time require servicing for the accumulation of data. When servicing is required, interrupts put the program in foreground. In between requests for servicing, the program returns to background for the performance of tasks in progress. Real-time programming is considerably more complicated than programming when only background tasks are being executed because at any one time one background and any number of foreground tasks may be in progress. Extreme care is required if the foreground tasks are not to interfere with each other or the background task.

The data acquisition section of the program contains the code which sets up various data acquisition sequences, services foreground data acquisition tasks, and handles the special sequence of operations which take place when data acquisition is completed. Since both foreground and background tasks are performed, two rather different

0 0 0 0 4 3 0 8 0 3 1

structures are needed. Background tasks are performed by data acquisition routines which are of a rather conventional form. Foreground tasks are performed by specially structured subroutines called drivers. These subroutines initiate the required action, service relevant interrupts while the action is in progress, and handles termination of the action. We can see how a typical driver works by considering the sequence of operations involved in the acquisition of a single interferogram.

Suppose we wish to measure an interferogram that is 150 points long. Each point is to consist of the average of 1000 measurements taken at one millisecond intervals for a total integration time of one second per point. The mirror is advanced after the measurement of each point is finished. After the mirror advance is completed, the system waits 210 milliseconds to allow the lock-in to settle before acquiring the next point. At the end of the run, the mirror is reset to its initial position with backlash removed, the data are punched on paper tape for permanent storage, and the operator is alerted with a teletype message. The command which initiates this sequence of operations is R, 150, 1, 1000, 210. The monitor passes control to the RUN routine which performs a number of tasks to prepare for the run. It then calls the driver .POINT to handle the acquisition of the first data point. After setting up timing, .POINT returns control to the monitor, which reenters command mode. At this time the system is ready to accept additional commands. Every millisecond an interrupt occurs. In foreground, .POINT calls for an ADC conversion, then returns to background. When

the conversion is finished a second interrupt occurs. In foreground, the result is added to an accumulator. The program returns to background. The 1001st timing interrupt indicates the end of the acquisition period. In foreground the contents of the accumulator are divided by 1000 and stored. A mirror advance is called for and the program returns to background. The interrupt signalling the completion of the mirror advance is serviced by calling for a 210 ms wait period. The interrupt at the end of the wait period indicates completion of the measurement of that point. The next point is measured by calling .POINT, starting the whole sequence again. After the last point has been acquired, special code is issued to perform the termination operations described above. The acquisition of each point in the above example required 2003 interrupts. Since the average interrupt service routine required about 15 μs, the system was in foreground mode for only about 30 ms during that time. Therefore only about 2% of the total measurement time was used for data acquisition. The rest of the time was available for background task execution.

The concepts used in writing the background routines and drivers for I/O operations are essentially the same as the concepts used for data acquisition programming. One interesting complication is due to the teletype. Typing on the teletype keyboard does not cause character to be printed. The characters are only transferred to the computer. Usually the keyboard driver transmits these characters to the teletype

0 0 0 4 3 0 8 0 3 2

printer as they are received, "echoing" the typed information; however, this introduces a possible source of conflict between printing a message on the teletype and echoing a typed character. The teletype keyboard and printer drivers were specially designed with this conflict in mind.

The bulk of the data analysis needed for FTS is performed on-line by special data analysis subrountines which are capable of performing their calculations on typical data records (containing perhaps 1000 data points) in a second or two. Although detailed discussion of the algorithms will be postponed to Appendix D, it is appropriate to indicate here some of the data analysis options that are available in FTS-11. The most important data analysis task is the transformation of data from an interferogram in path difference space to a spectrum in frequency space by means of a Fourier transform. A typical interferogram consists of a few points before zero path and a large number of points after zero path. The data is usually offset by a baseline. The record is prepared for transformation by subtraction of the baseline and determination of the zero path point. The transformation is then performed, starting at zero path, using a fast Fourier transform (FFT) algorithm based on a modification[3] of the well-known Cooley-Tukey algorithm.[4] Finally, the real and imaginary parts of the result are separated for ease of display, since only the real parts are normally of interest. The software will prepare, transform, and sort a 1024 point interferogram in about one second.

A typical source of error occurs when the interferogram is not properly sampled. Although the sampling occurs at uniformly spaced intervals, it may not happen that one of the sample positions is precisely at zero path. This type of error and certain other types which may occur can be corrected by a procedure known as phase correction. We have implimented a phase correction algorithm due to Forman, Vanasse, and coworkers.[5] Other operations which are used to help prepare spectra for interpretation are apodization and averaging of several interferograms in order to improve the signal-to-noise ratio.

In FTS, the spectrum obtained from the interferogram is the product of the sample spectrum and an intrumental response function. In order to remove the instrumental response, the spectrum must be divided by a background spectrum. The FTS-11 program can calculate the ratio of two spectra, including the effect of amplifier gain differences between numerator and denominator. It is also able to calculate the negative of the natural logarithm of the resultant ratio. This is useful when the sample property being studied may be characterized by an absorption coefficient. The negative natural logarithm is proportional to the absorption coefficient.

Although in a minimal system only one interferogram or spectrum need be stored in computer memory at a time, it is very convenient to be able to store several records at once. The management and control of the various storage areas or buffers is the job of the buffer management section of the program. A directory is kept of all the buffers implemented in a given version of the program. The directory

0 0 0 4 3 0 8 0 3 3

entry for each buffer contains the address of its first word, its length, the number of words currently being used, and a flag. This information is used to locate the buffer and control its use. Since there is only a finite amount of memory available for buffer storage, the user must choose between having a large number of short buffers and a smaller number of longer ones. This choice will depend on the requirements of the experiment being performed. The use of a directory makes it very simple to develop different versions of the program which implement different choices of buffer allocation. In principle, the buffers could be allocated dynamically at the time the program is loaded into the computer, but this has not been implemented.

The final section of the program consists of the utilities. These are system subroutines used by the rest of the program to perform particular tasks. Examples of utilities are routines to save and restore working registers needed by both foreground and background tasks, error handling routines, code conversion routines, and floating point arithmetic routines. Code conversion is used to convert between the binary information stored in the computer and the character-string representations of that information used by I/O devices such as the teletype. Floating point software is required because the computer is designed to perform integer arithmetic. In order to handle non-integer quantities, a floating point format (conceptually similar to scientific notation) is used. Special subroutines are necessary to perform arithmetics operations using this format.

CHAPTER III REFERENCES

1. Although in most computers on-line (core and semiconductor) memory is not considered a peripheral device and uses a special data path (memory bus), in all but the fastest members of the PDP-11 series, there is no memory bus. Communication between memory and the central processor utilizes the Unibus.
2. Actually this is an oversimplification. Any device except memory may become the controller of the Unibus (Bus Master) and any device except the processor may be the slave device. Read operations are from the slave to the master; write operations are from the master to the slave. Bus operations exist to transfer mastership from one device to another.
3. W. M. Gentleman and G. Sande, Fall Joint Computer Conference Proceedings 29, 563 (1966).
4. J. W. Cooley and J. W. Tukey, Math. of Comput. 19, 297 (1965).
5. M. L. Forman, W. H. Steel, and G. A. Vanasse, J. Opt. Soc. Am. 56, 59 (1966); H. Sakai, G. A. Vanasse, and M. L. Forman, J. Opt. Soc. Am. 58, 84 (1968).

0 0 0 0 4 3 0 8 0 3 4

IV. FAR INFRARED EPR OF $Al_2O_3:Mn^{3+}$

One of the subjects of continuing interest in the study of dilute transition metal ions in crystals has been systems with orbitally degenerate ground states, in which the Jahn-Teller (JT) effect plays an important role. An example is the ^5E ground state of the isoelectronic species Mn^{3+} and Cr^{2+} in Al_2O_3. Previous experimental work on these ions has principally involved measurements of thermal conductivity[1,2] and acoustic paramagnetic resonance.[3-6] Some preliminary infrared measurements have also been reported.[7,8] In this chapter I will discuss results of a detailed study of the far infrared spectrum of Mn^{3+} in Al_2O_3.[9] In the next chapter I will discuss measurements of γ-irradiated ruby, which is expected to contain Cr^{2+}.

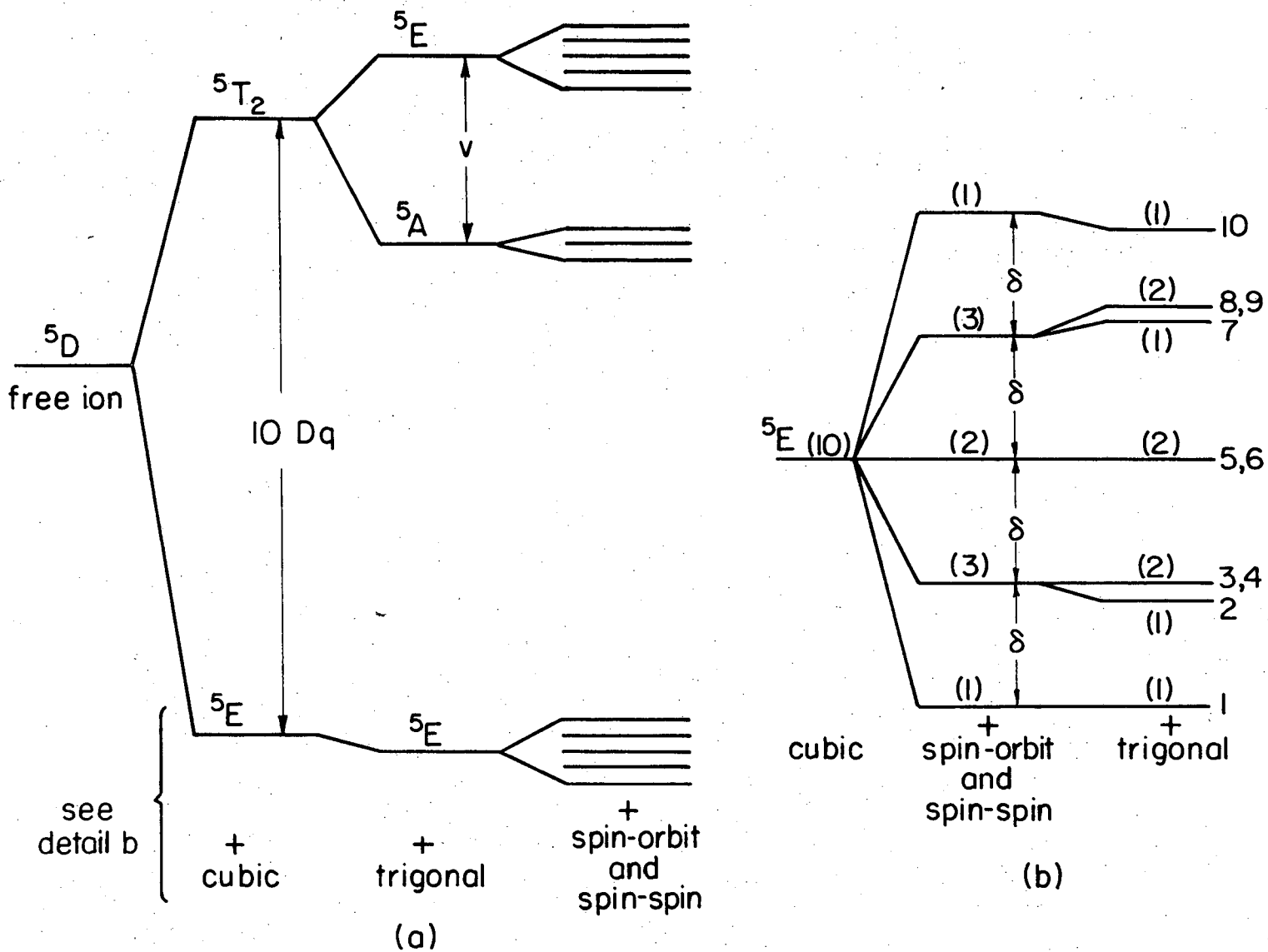
In order to describe our data we have developed a phenomenological model for the low lying states of Mn^{3+} in Al_2O_3 based on a crystal field Hamiltonian in the absence of the JT interaction. The strengths of the crystal field terms are adjusted to fit the data. We find that various parameters have smaller values than anticipated for a purely electronic Hamiltonian. We interpret this quenching as due to the effect of the JT interaction. An overview of the model may be obtained using simple group-theoretic arguments.[10] The model will be described in detail later.

The Mn^{3+} ion is believed to enter Al_2O_3 at substitutional Al^{3+} sites which have site symmetry C_3. The crystal field at such a site is predominantly cubic with a trigonal distortion. As is shown in

Fig. 13(a), the free ion $(3d^4)\,^5D$ Hund's rule ground state of Mn^{3+} is split by the cubic field into $^5E + {}^5T_2$ with ^5E lower. The trigonal field splits the $^5T_2$ state, but leaves the ^5E level degenerate. This degeneracy is lifted by the spin-orbit interaction in second order, and by the spin-spin interaction in first order.

In Fig. 13(b) we show the splitting of the ^5E ground state in more detail. Early crystal field calculations[3] which did not allow for the quenching of the trigonal field predicted that the lowest lying ^5E level was a doublet. Following a suggestion by C. A. Bates we consider a strongly quenched trigonal field. Applying first the spin-orbit and spin-spin interactions to the ^5E level of Mn^{3+} in a cubic field, five equally spaced levels with separation δ result. The lowest level is a singlet. In order of increasing energy the other levels are a triplet, a doublet, a triplet, and a singlet. The trigonal field splits each triplet into a doublet and a singlet and slightly changes the level splittings. If the trigonal field were not quenched, the singlet split off from the lowest triplet would be degenerate with the ground singlet. Our data show that the lowest level of the system is in fact a singlet. For purposes of discussion it will be convenient to label the energy levels as is indicated by the numbers on the right hand side of Fig. 13(b). This is the order of the levels in the absence of an applied magnetic field.

0 0 0 0 4 3 0 8 0 3 5



XBL752-5767

Fig. 13. $Al_2O_3:Mn^{3+}$ Schematic Level Diagram.

A. Experimental Details

Our results were obtained by measuring the transmission of samples using Fourier transform spectroscopy. Most of the measurements were made using the PMI. The interferometer output was coupled, using evacuated 1.1 cm light pipes, to the sample cryostat shown in Fig. 14. It consisted of a 4 1/2 inch diameter main helium vessel surrounded by a vacuum jacket and a liquid nitrogen shield, not shown in the figure. The insert consisted of a 1 1/8 inch diameter sample chamber, NbZr superconducting solenoid capable of producing fields to 55 kG, and a doped-Ge bolometric detector mounted in a separate vacuum chamber. Sample temperature could be varied below 4.2 K by pumping the main helium bath. In order to elevate the sample temperature above 4.2 K, the vent ports were sealed and the sample chamber was filled with exchange gas. A heater controlled by a standard temperature regulator maintained the desired temperature. This apparatus, which has been described in detail elsewhere,[11,12] has the desirable feature that samples could be changed during a run with only a minimal loss of helium. The bolometer was heat sunk to a separate helium pot which could be easily maintained at 1.1 K.

The samples were wrapped in thin Al foil to minimize the radiation losses out the sides and thus form a low Q transmission cavity. Radiation was coupled into and out of the sample, which was placed in the center of the solenoid, using conical light concentrators. The bolometer detected the radiation transmitted by the sample. The

0 0 0 0 4 3 0 8 0 3 6

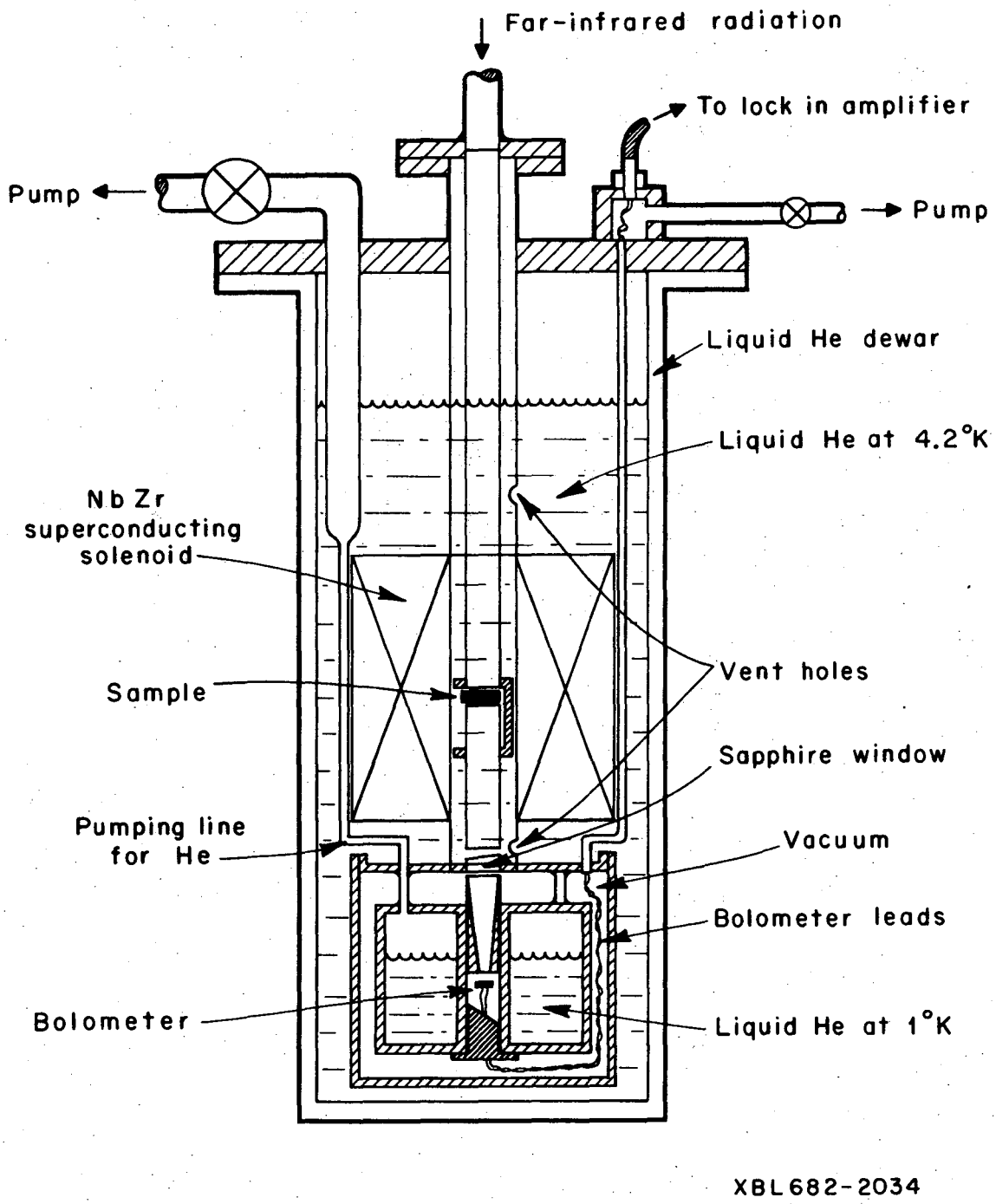


Fig. 14. Cryostat

computerized data acquisition techniques used in these measurements were described above. The instrumental response function was removed by computing ratios of pairs of spectra as is shown in Fig. 15. The peak at 9.35 cm^{-1} is due to an absorption line in the zero field denominator. Other features are due to absorption lines in the numerator, measured in an applied field of 30 kG. The dashed lines are the one standard deviation error bars and indicate the signal-to-noise ratio obtained in these measurements. The poor signal to noise ratio at either end of the spectrum is due to the division of small numbers, since the spectral intensity approaches zero in these regions.

The four samples used in this work were obtained from the sources indicated in Table 3. Spectrochemical analyses were made in order to identify the impurity responsible for the observed absorption lines. Since the properties of many ions in Al_2O_3 have been characterized, most of the ions which could be responsible for the observed absorption could be eliminated. The most likely candidates were Mn^{3+}, Cr^{2+}, Cr^{4+}, and Fe^{2+}. By comparing the observed line strengths in different samples with the respective spectrochemical analyses, it is possible to rule out all likely sources of the lines, except Mn^{3+}.

B. Experimental Data

We have measured the transmission of $Al_2O_3:Mn^{3+}$ in the frequency range from 3 to 30 cm^{-1} as a function of applied magnetic fields up to 55 kG. The points in Figs. 16 and 17 are the positions of observed far infrared absorption lines plotted as a function of applied magnetic field. The data were separated into ground state transitions (Fig. 16)

0 0 0 4 3 0 8 0 3 7

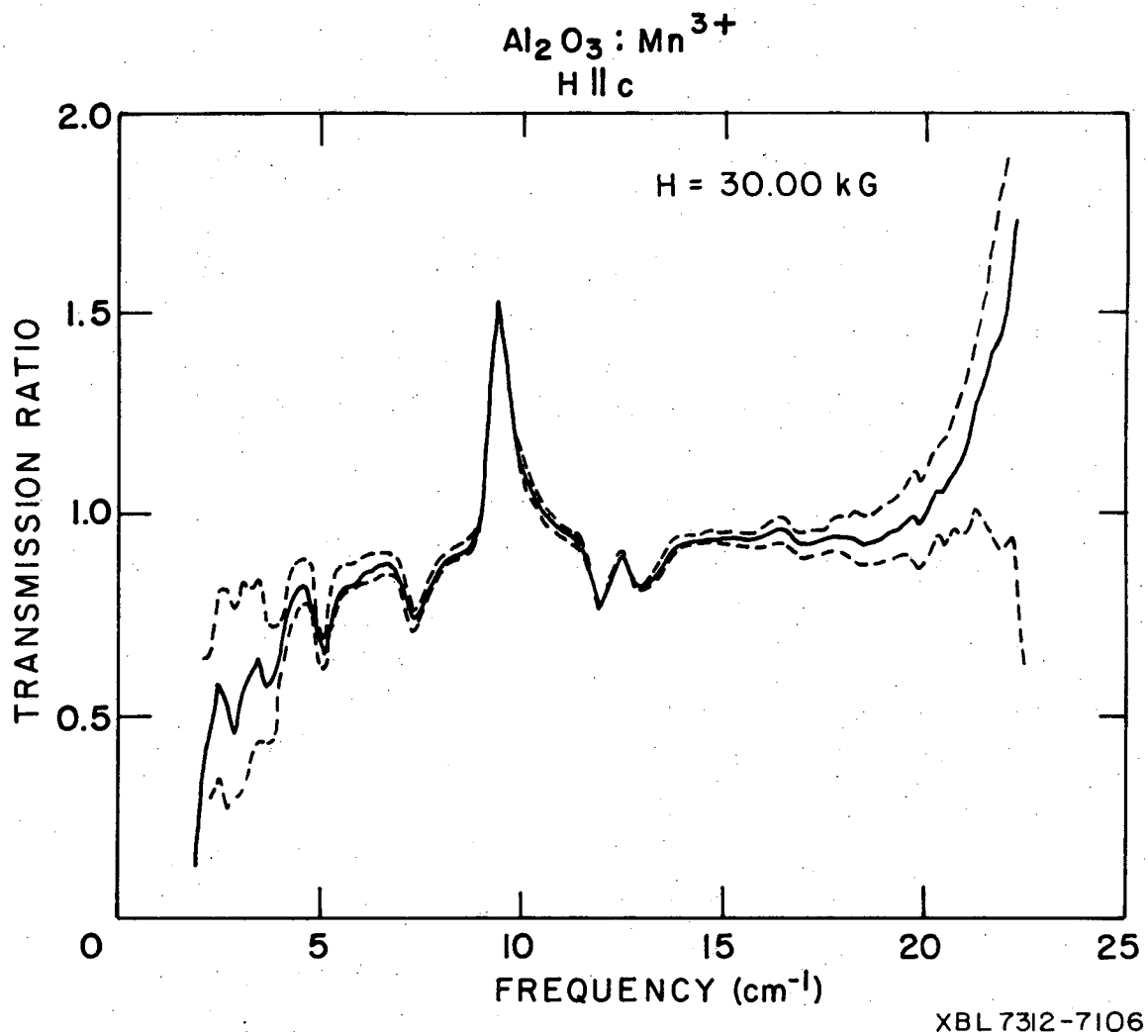
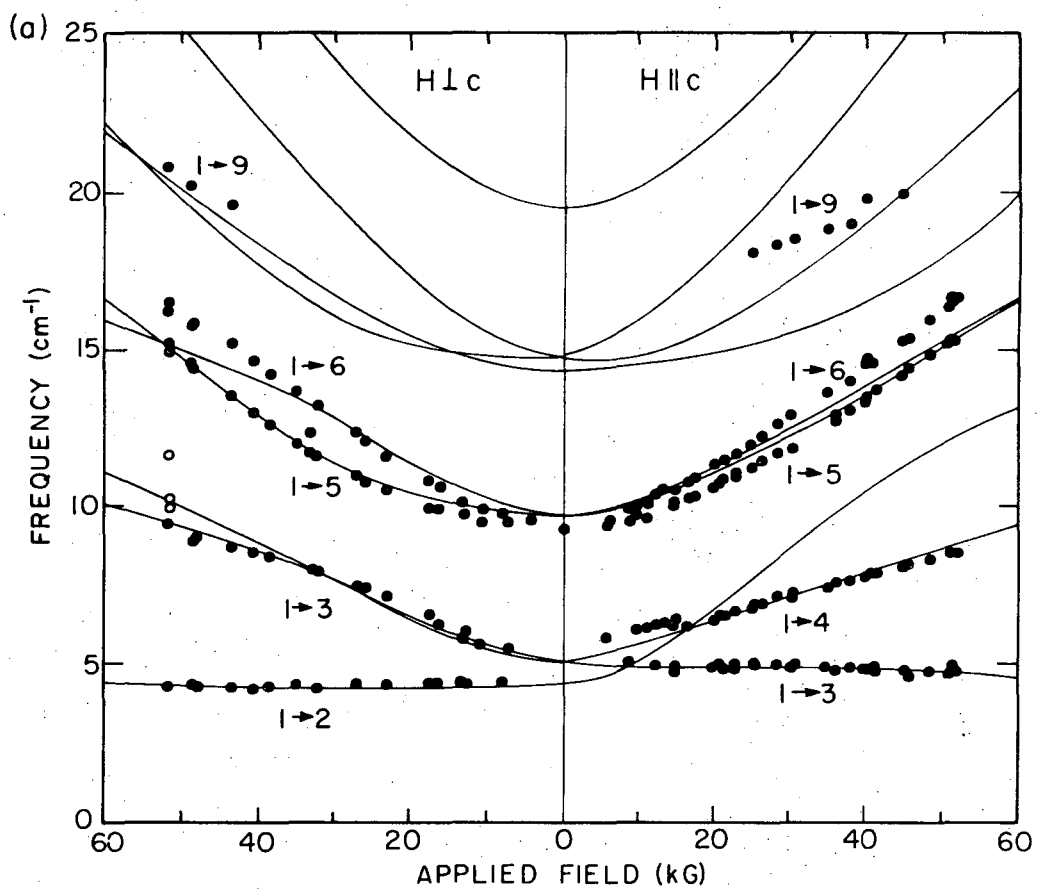


Fig. 15. Field On--Field Off Ratio.

Table 3. Sample Data.

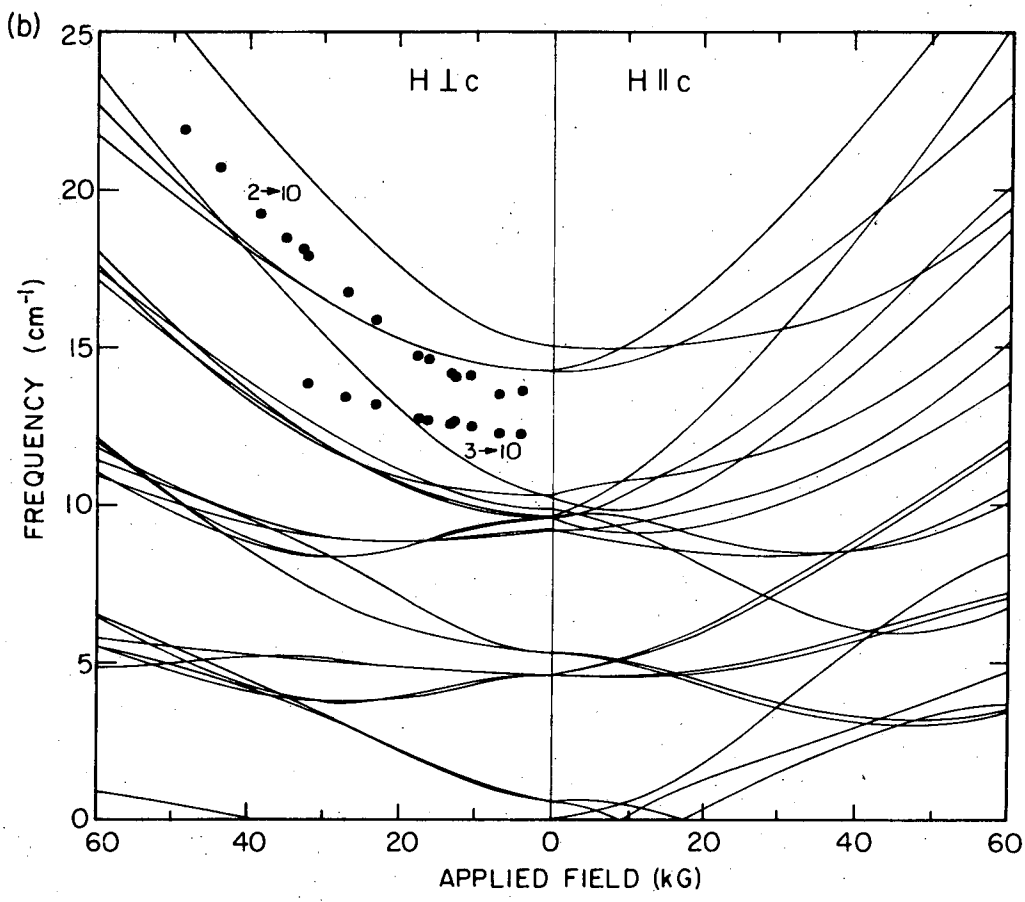
Sample	Source	Length (cm)	Mean Diameter (cm)	℄	Analysis (% by weight) Mn	Cr	Fe
S-4	A. C. Pastore, Korad Corp.	1.5	0.8	∥	.01	.5	.01
S-5	W. Brewer, National Bureau of Standards	1.9	0.3	∥	.001	<.001	.001
S-14	Hrand Djevahirdjian S.A. (Switzerland)	2.5	1.0	∥	.02	<.001	.1
S-18	C. Sahagian, Air Force Cambridge Research Lab.	3.9	0.8	⊥	.02	<.001	<.001

0 0 0 0 4 3 0 8 0 3 8



XBL 751-5438A

Fig. 16. Ground State Transitions.



XBL 751-5438B

Fig. 17. Excited State Transitions.

0 0 0 0 4 3 0 8 0 3 9

and excited state transitions (Fig. 17) by measurement of the temperature dependence of the line strengths. The data for which the applied field was oriented parallel to the crystallographic c axis were measured using samples S-4, S-5, and S-14. The data for $H \perp c$ were measured using S-18.

Two absorption lines are observed to intersect zero field near 5 cm^{-1} in both the $H \parallel c$ and $H \perp c$ orientations. In the parallel field case, the lines are degenerate at $H = 0$ with a frequency of 5.2 cm^{-1}. In the perpendicular field case the modes intersect $H = 0$ at 4.5 cm^{-1} and 5.2 cm^{-1}. Temperature dependence establishes that all of these modes are due to transitions from a singlet ground state. Furthermore, the temperature dependence data show that the state which lies 4.5 cm^{-1} above the ground state is the lowest lying excited state of the system. The transition to this state is therefore labeled $1 \rightarrow 2$. The two degenerate states which lie 5.2 cm^{-1} above the ground state in zero field are states 3 and 4. The $H \parallel c$ transitions are therefore labeled $1 \rightarrow 3$ and $1 \rightarrow 4$. The $H \perp c$ absorption line which intersects $H = 0$ at 5.2 cm^{-1} is either a $1 \rightarrow 3$ or a $1 \rightarrow 4$ transition. Comparison of model calculations based on these two possible assignments indicate that the transition is probably $1 \rightarrow 3$. The intensities of these transitions are field-dependent, as indicated in Fig. 18. They have zero intensity at zero applied field, while at 50kG, these lines dominate the observed spectra.

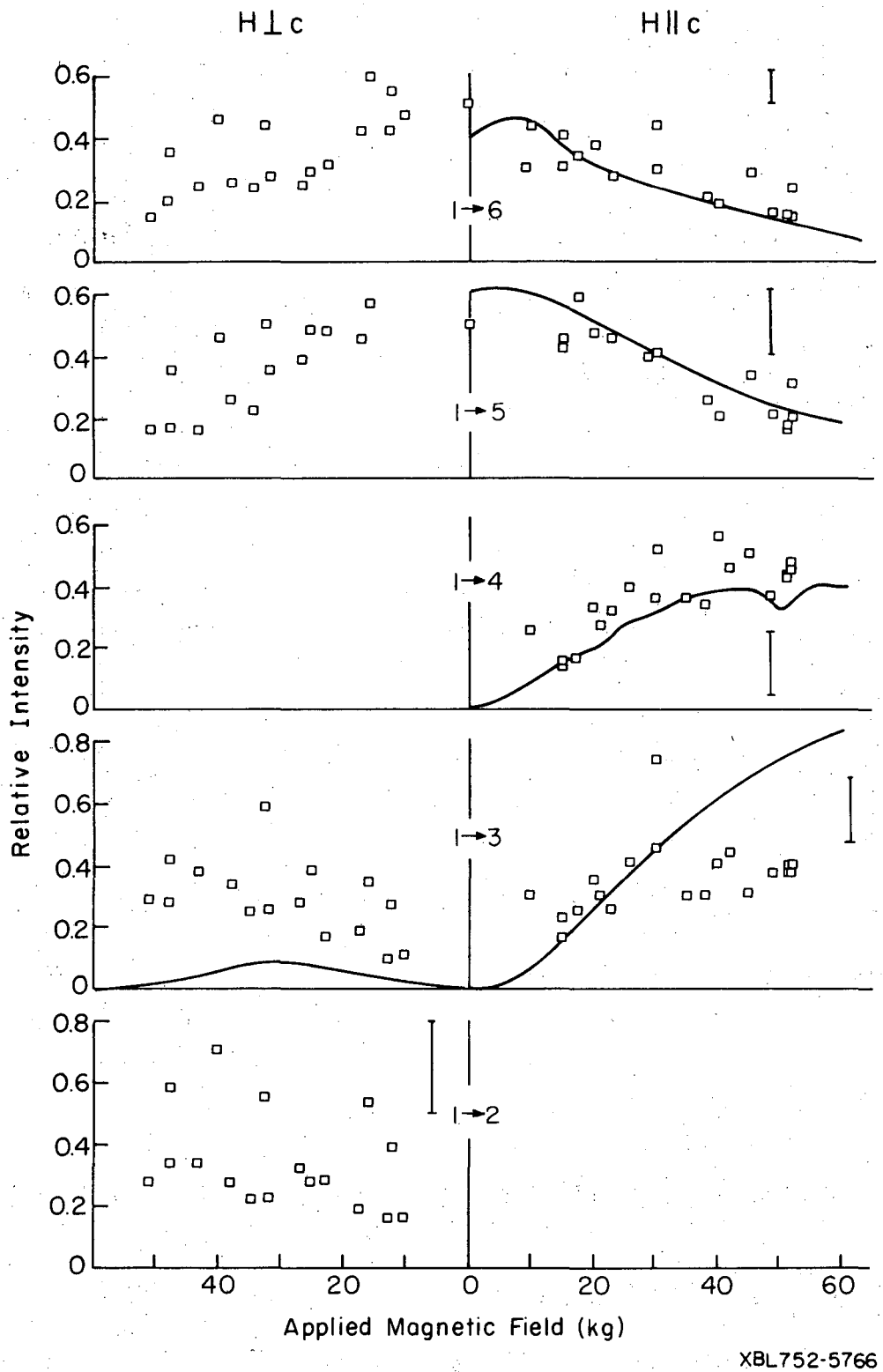


Fig. 18. Ground-State Transitions--Peak Absorption Coefficients.

0 0 0 0 4 3 0 8 0 4 0

Two absorption lines are observed which intersect H = 0 at 9.35 cm^{-1}. These lines are also ground state transitions and have been labeled $1 \rightarrow 5$ and $1 \rightarrow 6$. In zero field the degenerate $1 \rightarrow 5,6$ line is the dominant absorption feature in the spectra of all the samples studied. The intensity of these lines decreases with field, as is shown in Fig. 18. In all three of the samples used for the $H \parallel c$ studies the H = 0 line was asymmetric with a linewidth (full width at half peak absorption coefficient) of 0.6 to 0.8 cm^{-1}, depending on the sample. In S-18, used for the $H \perp c$ studies, the H = 0 line was symmetric with a linewidth of 0.3 cm^{-1}. The linewidths are not temperature dependent below 20K. We have not identified the mechanism responsible for broad lines in certain samples.

The weak ground state transition labeled $1 \rightarrow 9$ is observable only at the highest available fields. This transition was observed only in the two samples with the greatest Mn concentration.

In the $H \perp c$ orientation (sample S-18) the two weak excited state transitions shown in Fig. 17 were observed. These lines were only observed in the sample containing the largest total number of Mn ions. The separation between the two observed excited state transitions is equal to the separation between the $1 \rightarrow 2$ and $1 \rightarrow 3$ transitions at all values of applied field. Thus, it is reasonable to suppose that the excited state lines are due to transitions between the states 2 and 3 and some single higher-lying state. The best candidate for this higher-lying state is the singlet level 10. The lower excited state transition in Fig. 17 is, then, $3 \rightarrow 10$, while the higher transition

is 2 → 10. The temperature dependence of these lines seems to confirm this identification.

Radiation normally propagates down the sample chamber in a direction parallel to H. Although for thin samples mirrors can be installed which enable propagation through the sample perpendicular to H, our samples were too large for this to be done. Since in our measurements radiation propagated parallel to H, it was polarized perpendicular to H. Samples S-4, S-5, and S-14 are cylinders with the optic axis parallel to the cylinder axis. They could only be studied with radiation polarized $\underset{\sim}{E} \perp \underset{\sim}{c}$. Sample S-18, however, had the c axis perpendicular to the cylinder axis and could be oriented so that both $\underset{\sim}{E} \perp \underset{\sim}{c}$ and $\underset{\sim}{E} \parallel \underset{\sim}{c}$ polarizations could be studied. A linear polarizer consisting of 10^3 gold strips per inch on a Mylar substrate was attached to this sample with rubber cement. Typical transmission ratios are displayed for the two polarizations in Fig. 19. The spectra in the numerators were measured in an applied field of 23 kG while the spectra in the denominators were, as usual, zero field measurements. These spectra were all measured at 4.2 K. The dashed lines again indicate the error bars. The solid arrows in the $E \perp c$ ratio, Fig. 19(a) indicate observed absorption lines. The dashed arrows in the $E \parallel c$ ratio, Fig. 19(b), are drawn at the same frequencies as the solid arrows for $E \perp c$. These dashed arrows indicate where the $E \perp c$ modes would occur were the transitions allowed in this polarization. Some weak features are seen in the $E \perp c$ spectrum, however, these features may be explained by imperfect polarization arising from multiple reflections in the sample. To the accuracy of the

0 0 0 4 3 0 3 0 4 1

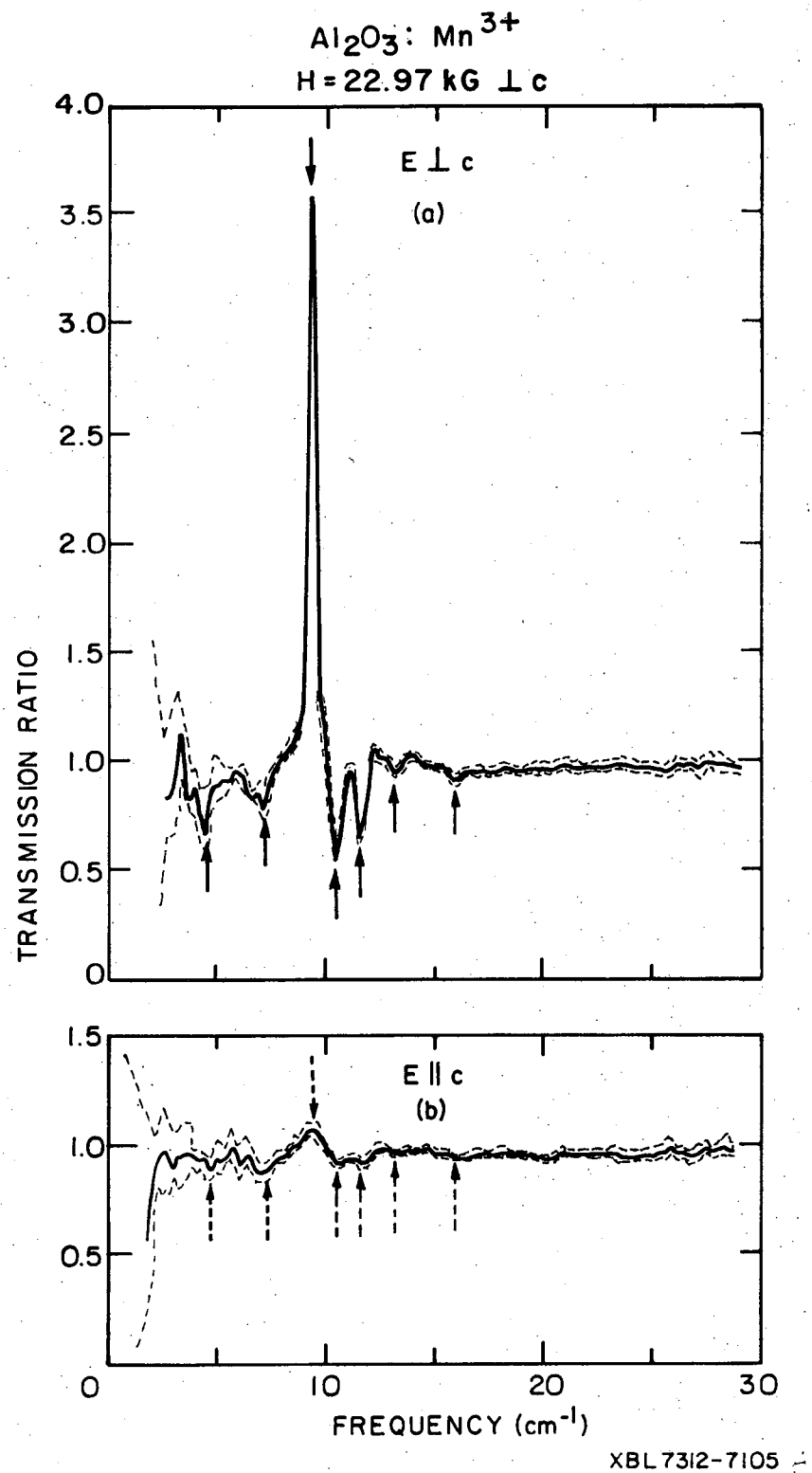


Fig. 19. Polarization Dependence.

experiment, these absorption lines are completely polarized. The three experimentally accessible combinations of orientation and polarization--$H \parallel c$, $E \perp c$; $H \perp c$, $E \perp c$; and $H \perp c$, $E \parallel c$--are sufficient to unambiguously assign the observed absorption lines to electric dipole transitions.

C. Theory

The JT effect is expected to play an important role in the states which we have studied. Several attempts have been made to calculate the lowest lying vibronic states of Mn^{3+} and the isoelectronic species Cr^{2+} in Al_2O_3.[13-15] Because of the extreme complexity of the problem, however, these efforts have met with limited success. Rather than attempt to calculate vibronic energy levels, we have adopted a phenomenological philosophy. Were there no JT effect operating, we would expect the system to be well described by a purely electronic crystal field Hamiltonian. By allowing the interaction strengths of various terms in such a Hamiltonian to be adjustable parameters, we can ask how well such a simple Hamiltonian is capable of fitting our data. We find that the fit is surprisingly good. The values of the interaction strengths which give the best fit to the data can be compared to the values of these terms one would expect without a JT interaction. We believe that the deviations of these parameters from their expected (no JT effect) values to be related, although perhaps not directly, to the Ham quenching factors[16] of a more complete vibronic theory.

0 0 0 0 4 3 0 8 0 4 2

Our model Hamiltonian is the sum of crystal field, spin-orbit, spin-spin, and Zeeman terms. Since hyperfine splittings are too small to be resolved in our experiments, we neglect them.

$$\mathcal{H} = V_{xtal} + V_{so} + V_{ss} + V_z \quad . \tag{55}$$

The crystal field potential V_{xtal} consists of a cubic term, a trigonal distortion, and an inversion-breaking term. The cubic and trigonal fields are written using the usual spin operators,[17]

$$\begin{aligned} V_{cubic} &= -\frac{2}{3} B_4 (O_4^o + 20\sqrt{2}\, O_4^3) \ , \\ V_{trig} &= B_2^o O_2^o + B_4^o O_4^o \ , \end{aligned} \tag{56}$$

where the B coefficients are related to McClure's parameters[18]

$$\begin{aligned} Dq &= 12\, B_4 - (\tfrac{14}{3}) B_4^o \\ v &= 9\, B_2^o + 80 B_4^o \\ v' &= \sqrt{2}\, (3B_2^o + 20B_4^o) \ . \end{aligned} \tag{57}$$

A Hamiltonian consisting only of cubic and trigonal fields has D_{3d} symmetry. In order to correctly represent the splitting of the 9.35 cm^{-1} line in an applied field, the symmetry of our Hamiltonian must be no higher than C_3. We reduce the symmetry by adding an inversion-breaking term V_{ib}. Odd order terms can occur in the Hamiltonian in the absence of inversion symmetry. Since these terms vanish within the ^5D manifold (L = 2, S = 2), they admix excited configurations into the ^5D state.

It is important to consider such admistures when computing transition probabilities, but they may be neglected when calculating the energies of the low lying levels of the system. The only symmetry-allowed even-order term not already included in the crystal field is

$$A_4^3 \propto \frac{1}{2i}\,(Y_4^3 + Y_4^{-3}) \quad . \tag{58}$$

The dominant inversion-breaking term is, then,

$$V_{ib} = B_4^i A_4^3 \quad . \tag{59}$$

Within the ^5D manifold, the spin-orbit interaction may be written $V_{so} = \lambda \underset{\sim}{L} \cdot \underset{\sim}{S}$. The V_{ss} term,

$$V_{ss} = -\,\rho(\underset{\sim}{L} \cdot \underset{\sim}{S})^2 \ , \tag{60}$$

contains contributions from two sources. The first is the usual spin-spin interaction which may be written in an L-S coupling scheme as[17]

$$-\,\rho\{(\underset{\sim}{L} \cdot \underset{\sim}{S})^2 + \frac{1}{2}\,(\underset{\sim}{L} \cdot \underset{\sim}{S}) - \frac{1}{3}\,L(L+1)S(S+1)\} \ .$$

The second source is the admixture of excited terms (in particular, the ^3H term) of the $3d^4$ configuration through spin-orbit coupling. A typical contribution might be schematically written in the form

$$\sum_{^3H} \sum_{\text{states}} \frac{\langle {}^5D|\underset{\sim}{L} \cdot \underset{\sim}{S}|{}^3H\rangle\,\langle {}^3H|V_{xtal}|{}^3H\rangle\,\langle {}^3H|\underset{\sim}{L} \cdot \underset{\sim}{S}|{}^5D\rangle}{(E_{^5D} - E_{^3H})^2} \ .$$

0 0 0 4 3 0 8 0 4 3

Since these contributions are small, the approximation of including them in Eq. (60) seems justified. The Zeeman interaction is written $(k\underset{\sim}{L} + 2\underset{\sim}{S}) \cdot \mu_B H$ where k is included to allow for covalency effects. A typical value of k is 0.8 for transition metal ions in Al_2O_3.

The Hamiltonian defined above must be evaluated using four-electron wavefunctions appropriate to the ^5D free ion ground state. These wavefunctions are linear combinations of products of one-electron d-state wavefunctions. We first form Slater determinant wavefunctions which will be denoted by $(\ell_1^{s_1}\ \ell_2^{s_2}\ \ell_3^{s_3}\ \ell_4^{s_4})$. This wavefunction is a determinant of four one-electron wavefunctions, the first with quantum numbers $\ell_z = \ell_1$, $s_z = s_1$, the second with $\ell_z = \ell_2$, $s_z = s_2$, etc. Interchange of any two electrons changes the sign of the wavefunction.

$$(\ell_1^{s_1}\ \ell_2^{s_2}\ \ell_4^{s_4}\ \ell_3^{s_3}) = -(\ell_1^{s_1}\ \ell_2^{s_2}\ \ell_3^{s_3}\ \ell_4^{s_4}) \ .$$

Permutations are performed until the one-electron wavefunctions are listed in order of decreasing ℓ_z and decreasing s_z (consistent with the order of the ℓ's). Next we form the wavefunctions $|L_z, S_z\rangle$ appropriate to describe the ^5D free ion state by forming linear combinations of Slater determinants. The ^5D state wavefunction $|0,0\rangle$ may be written[19]

$$|0,0\rangle = -(6)^{-\frac{1}{2}}\{(2^+1^+-1^--2^-) + (2^-1^--1^+-2^+) + (2^+1^--1^+-2^-)$$

$$+ (2^-1^+-1^--2^+) + (2^-1^+-1^+-2^-) + (2^+1^--1^--2^+)\} \ . \qquad (61)$$

The other twenty-four wavefunctions are formed by applying the appropriate ladder operators to this state.

In order to evaluate the matrix elements of the Hamiltonian in the ^5D basis, we first must calculate the one-electron matrix elements. The one-electron matrix elements for the spin-orbit, spin-spin, and Zeeman interactions can be calculated directly. The crystal field matrix elements are more difficult. Since the spin operators O_ℓ^m transform according to representations of the rotation group, the Wigner-Eckhart theorem may be used to help calculate their matrix elements. For example, O_4^3 transforms according to $(Y_{43}-Y_{4-3})$ where the Y's are spherical harmonics. The matrix elements of Y_{LM} between angular momentum eigenstates is given by

$$\langle \ell'm' | Y_{LM} | \ell_m \rangle = (-1)^{-m'} \left[\frac{(2\ell'+1)(2L+1)(2\ell+1)}{4\pi} \right]^{1/2}$$

$$\times \begin{pmatrix} \ell' & L & \ell \\ -m' & M & m \end{pmatrix} \begin{pmatrix} \ell' & L & \ell \\ 0 & 0 & 0 \end{pmatrix} , \qquad (62)$$

using the conventional notation for Wigner 3-j symbols.[20] By the Wigner-Eckhart theorem, the matrix element $\langle \ell'm' | O_4^3 | \ell m \rangle$ must be proportional to the matrix element $\langle \ell'm' | (Y_{43}-Y_{4-3}) | \ell_m \rangle$. Using compiled tables of 3-j symbols,[20] these matrix elements are easily calculated. Radial integrals are incorporated in the adjustable parameters. Matrix elements of the Slater determinant wavefunctions are related to these one-electron wavefunctions by well-known rules.[21]

0 0 0 4 3 0 8 0 4 4

Matrix elements between ^5D states are calculated from the Slater determinant matrix elements.

The 25-dimensional Hamiltonian matrix was calculated and a subroutine written to evaluate the matrix for arbitrary values of B_4, B_2^o, B_4^o, B_4^i, λ, ρ, and k. The matrix was diagonalized using standard FORTRAN procedures and the schematic level diagram of Fig. 13 was verified by studying the general properties of the Hamiltonian as a function of the parameters. The splittings δ shown in Fig. 13(b) are given by[17]

$$\delta = 6(\rho + \frac{\lambda^2}{10Dq}) \ . \tag{63}$$

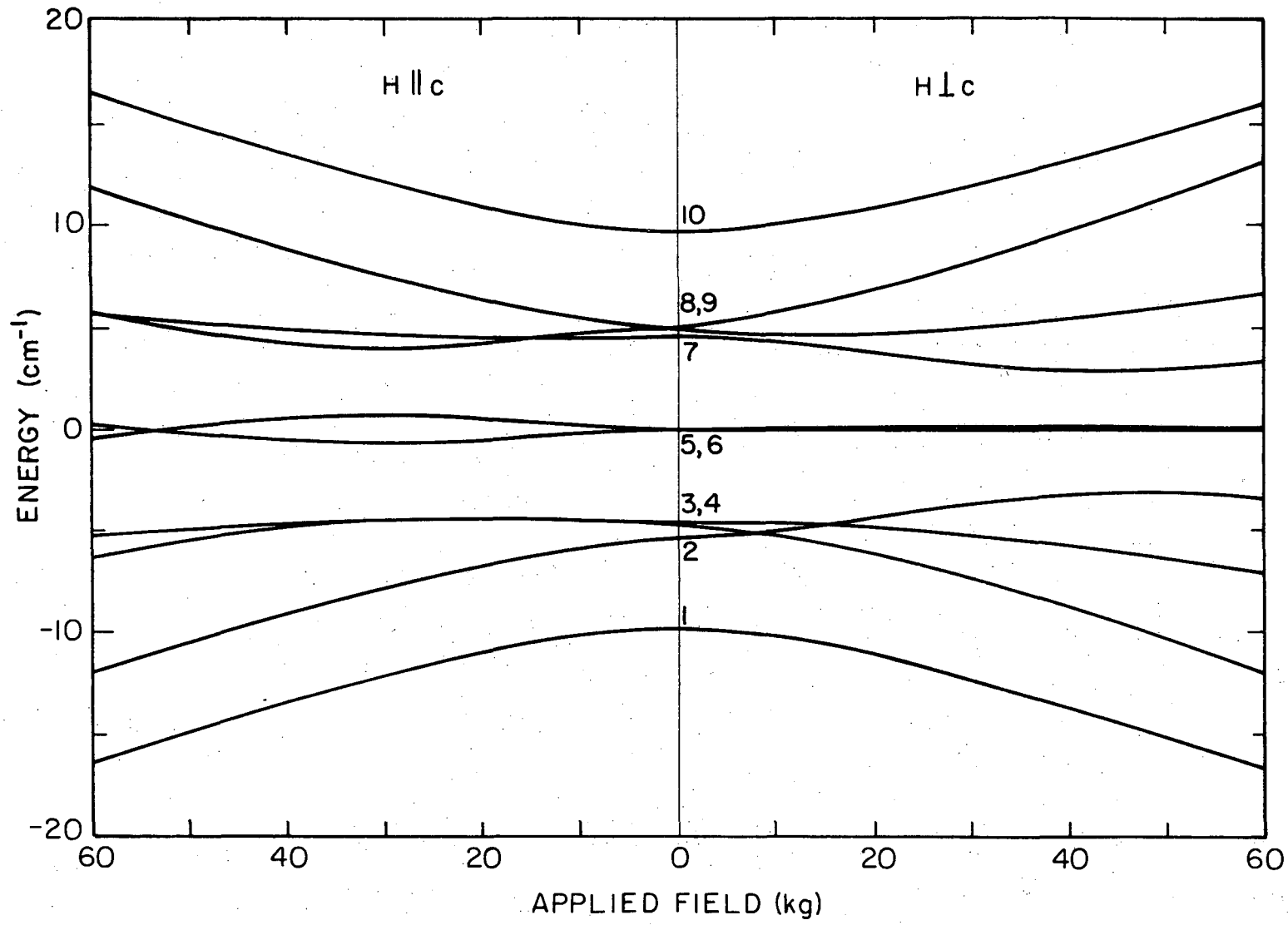
The inversion breaking parameter B_4^i has little effect on the zero field energy levels. With a finite magnetic field, it split 5 and 6, as expected. It also has large effects on the field dependence of the other energy levels. The orbital Zeeman effect is completely quenched within ^5E, the parameter k having no effect on the field dependence of the energy levels.

The adjustable parameters were determined by fitting the field-frequency data of Figs. 16 and 17 using a conventional least-squares procedure. It was found that good fits could not be achieved simultaneously for the ground and excited state data. A fit was attempted using the ground state transitions alone, but there were more adjustable parameters than could be determined from the data. In order to obtain a fit, B_4 was fixed. This parameter, which measures the cubic field, is not expected to be quenched by a JT interaction and can be determined from optical data. Unfortunately, the value of B_4

is uncertain. Measurements by McClure[18] and Forman[22] suggest values of 160.4 cm^{-1} and 142.1 cm^{-1} respectively. Fits were obtained using a number of values of B_4 between 140 and 165 cm^{-1}. Although the parameters obtained differed for each value of B_4, the resultant energy level diagrams were very similar. A typical energy level diagram is shown in Fig. 20. The parameters obtained using McClure's and Forman's values of B_4 are listed in Table 4. Considering the uncertainty in B_4, our parameters are determined only to within 15%. Some results, however, are clear. The trigonal field and spin-orbit interactions are strongly quenched. The inversion-breaking term is small, typically several cm^{-1}. The parameter ρ is about .8 cm^{-1}.

The three most important results of the model are the values of v, λ, and ρ. Because of JT quenching, the trigonal field parameter v has only 60-70% of its expected value. The spin-orbit coupling is very strongly quenched, having only about 10% of its free ion value. Finally, the parameter ρ has an unexpectedly important role in the structure of the ^5E state. Since λ is so strongly quenched, $\delta \approx 6\rho$. The splittings δ, however, are the most obvious feature of the ^5E energy level diagram. The parameter ρ is much larger than expected from a spin-spin interaction alone, suggesting that there is a considerable contribution to ρ from second-order spin-orbit coupling. This admixture of higher lying terms has a particularly important role in determining the structure of the ^5E state.

0 0 0 0 4 3 0 8 0 4 5



XBL752-5769

Fig. 20. Low Lying Energy Levels of $Al_2O_3:Mn^{3+}$--Theory.

Table 4. Interaction parameters for $Al_2O_3:Mn^{3+}$

	Dq	v	v'	B_4	B_2^o	B_4^o	B_4^i	λ	ρ	k
Without JT	1947[a]	1950	600	160.4	173.7	-4.837	-	87[c]	.18[d]	.8
	1727[b]			142.1						
Theory (e)	1946.4	1359.8	335.2	160.4	109.9	-4.634	-3.138	6.101	.8074	-
Theory (f)	1725.2	1234.8	298.6	142.1	99.0	-4.296	-2.805	6.059	.8069	-
Quenching (e)		.70	.56		.63	.96		.07		
Factor (f)		.63	.50		.57	.89		.07		

[a] McClure (Ref. 18).
[b] Forman (Ref. 22).
[c] Free ion value.
[d] Spin-spin interaction only.
[e] Calculation based on a value of 160.4 cm^{-1} for B_4.
[f] Calculation based on a value of 142.1 cm^{-1} for B_4.

0 0 0 0 4 3 0 8 0 4 6

D. Comparison with Experiment

The predicted ground state transitions are shown as solid lines on Fig. 16. The fits to the $1 \rightarrow 2$, $1 \rightarrow 3$, and $1 \rightarrow 4$ transitions are excellent. We regard the fits to the $1 \rightarrow 5$ and $1 \rightarrow 6$ transitions to be satisfactory considering the many approximations and simplifications made in the model. Predictions of the temperature dependence of the peak absorption coefficients of the ground state transitions were made using the theoretical level diagram and simple population arguments. The predicted temperature dependence agreed within experimental error limits with measurements made at seven temperatures between 1.3 and 20K.

In order to compare theoretical and experimental line strengths, an extension of our model is required. Since electric dipole matrix elements vanish within the ^5D manifold, we must include configurational mixing in order to account for our experimental results. The lowest lying configuration which can mix with the $3d^4$ configuration to give the required mixing is $3d^34p$. The $3d^34s$ configuration lies below $3d^34p$ but cannot contribute to the electric dipole transition probability because a one-electron $3d \rightarrow 4s$ transition is electric dipole forbidden. In order to admix $3d^34p$ wavefunctions with our $3d^4(^5D)$ states we ignore term separations within $3d^34p$ and assume that all of the $3d^34p$ states lie an energy Δ above the ^5D states. If the interaction which mixes the configurations is V_m, then to first order in perturbation theory the unnormalized admixed wavefunctions are of the form

$$|3d^4\rangle_* = |3d^4\rangle + \sum_{\substack{3d^34p \\ \text{states}}} \frac{\langle 3d^34p|V_m|3d^4\rangle}{\Delta} |3d^34p\rangle \quad . \qquad (64)$$

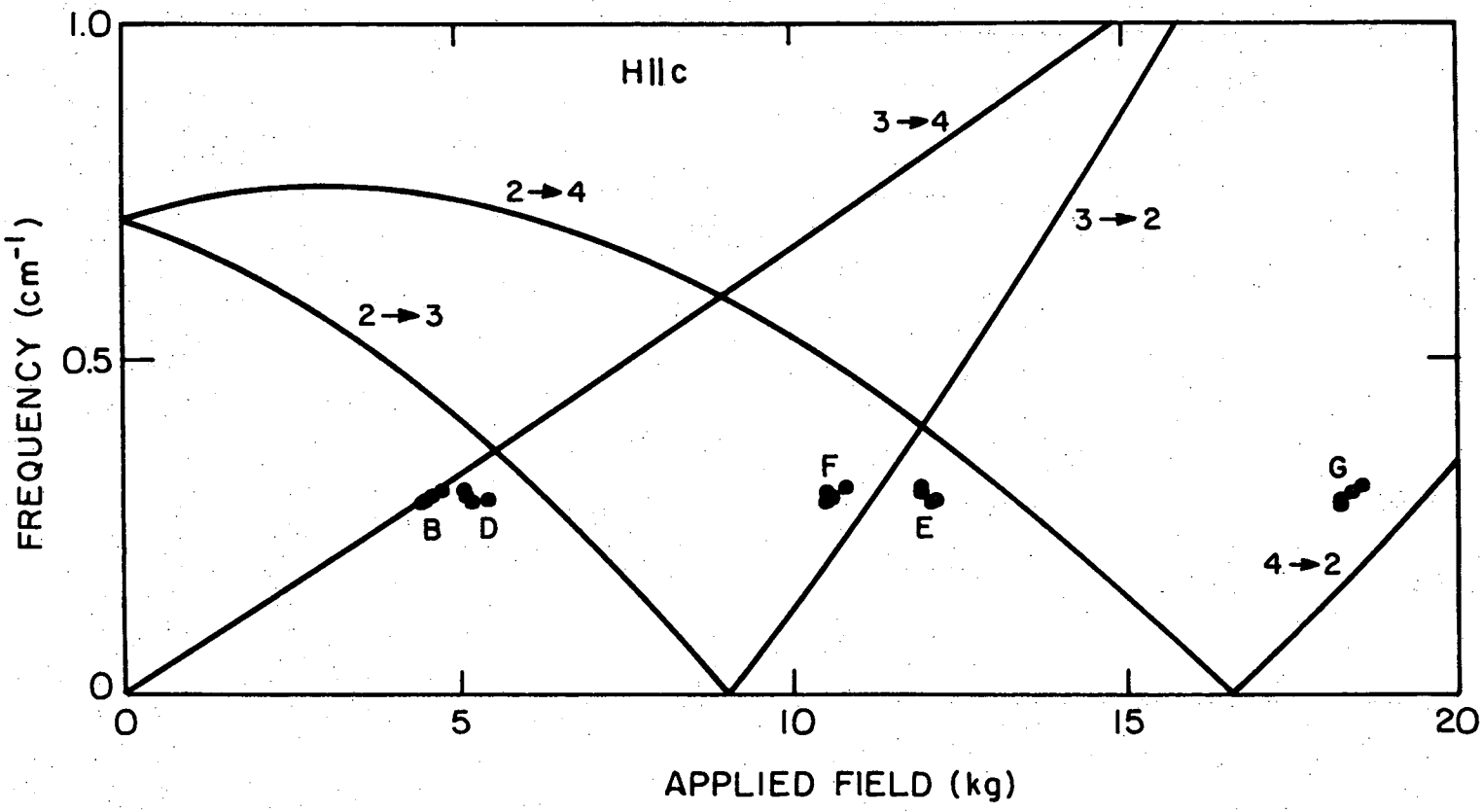
The mixing potential V_m is a sum over symmetry-allowed odd-order crystal field potentials. Since point charge model calculations indicate that the O_1^0 interaction is the largest allowed odd-order interaction for the atomic configuration of Al_2O_3, we assume that $V_m \propto O_1^0$, and calculate the admixed $3d^4$ wavefunctions. These wavefunctions are used to calculate the electric dipole matrix elements for the $^5D_{admixed}$ states, which in turn are used with the eigenfunctions of Eq. (55) and appropriate population factors to predict ground-state transition probabilities at 4.2K.

In order to compare calculated transition probabilities with observed peak absorption coefficients, a normalization factor is required because the concentration of Mn^{3+} in our samples is not known. Spectrochemical analysis yields an estimate of the total Mn concentration in the sample, but provides no information on the distribution of Mn valence states. Moreover we have not attempted to calculate either the strength of the O_1^0 interaction or Δ, so that an overall multiplicative factor for the theoretical line strengths is also unknown. We chose to normalize the data in Fig. 18 to the strength of the 9.35 cm^{-1} line in zero field. The equivalent normalization factor for the theoretical transition probabilities is the sum of the probabilities of the $1 \rightarrow 5$ and $1 \rightarrow 6$ transitions in zero field. The normalized transition probabilities for the lowest five ground state transitions are shown as the solid lines in Fig. 18. For $H \parallel c$ the model is seen to be in excellent agreement with experiment. For $H \perp c$, however, the model fails to account for the observed intensities. There could be

0 0 0 0 4 3 0 8 0 4 7

several reasons for the discrepancy. We have included only one of the odd-order operators that could cause configurational mixing and have neglected mixing with the excited terms of the $3d^4$ configuration which could also be important. One might also expect that the JT interaction, which has not been explicitly included in our model, would affect the transition probabilities in a significant way.

A further confirmation of the ability of our model to describe the low-lying states of Mn^{3+} may be obtained by considering the APR data of Anderson, Bates, and Jaussaud[5] (ABJ). In that paper, ABJ reported the observation of five APR lines in $Al_2O_3:Mn^{3+}$ at frequencies near 0.3 cm^{-1} which they labeled B, D, E, F, and G. These lines are due to transitions between states 2, 3 and 4. In Fig. 21, the solid lines are the theoretical transition frequencies for the transitions between these three states as a function of applied field. Since level 2 crosses levels 3 and 4 in this orientation, transitions $3 \rightarrow 2$ and $4 \rightarrow 2$ are plotted as well as $2 \rightarrow 3$, $2 \rightarrow 4$, and $3 \rightarrow 4$. The points are the data of ABJ. We identify the APR line B as being due to $3 \rightarrow 4$. D, E, F, and G are due to transitions $2 \rightarrow 3$, $2 \rightarrow 4$, $3 \rightarrow 2$, and $4 \rightarrow 2$ respectively. In this field orientation the observable far infrared transitions are $1 \rightarrow 3$ and $1 \rightarrow 4$. Transition $1 \rightarrow 2$ is forbidden. Hence, no information about the field dependence of level 2 in this field orientation was used to fit the model. The data of ABJ indicate, however, that the model correctly predicts the position of this level to within 0.1 cm^{-1}.



XBL 752-5768

Fig. 21. Comparison of Theory to APR Data for $Al_2O_3:Mn^{3+}$.

0 0 0 4 3 0 8 0 4 8

Although our model provides an excellent description of the ground state transitions, it is not as successful with the excited state transitions, 3 → 10 and 2 → 10. The theoretical excited state transitions arising from states 2, 3, and 4 are indicated as solid lines in Fig. 17. Clearly there are serious discrepancies between theory and experiment. Level 10 is located about 2 cm^{-1} lower than our model predicts, and its energy depends less strongly on applied field than the model indicates. In addition, at temperatures above 10K, the observed lines are much weaker than the predicted temperature dependence would indicate. It seems likely that a proper treatment of the JT interaction would reduce these discrepancies.

The simplest model of the JT interaction in Mn^{3+} is a cluster model in which the ^5E state interacts with an E symmetry vibrational state of the cluster. From group theoretic considerations

$$^5E_{electronic} \times E_{phonon} = {}^5E_{vib} + {}^5A_{vib} + {}^5A_{vib} , \qquad (65)$$

where the subscript 'vib' stands for vibronic states. The JT interaction removes the degeneracy of these vibronic states. Our data clearly indicate that the $^5E_{vib}$ state is lowest.

This simple JT model could account for some of the discrepancies between our theory and the experiment. Level repulsion from the $^5A_{vib}$ states could lower the energy of level 10 and affect its field-dependence, thus improving the poor fit for the observed excited state transitions. In addition, including ten energy levels (two $^5A_{vib}$ states)

at 25 cm^{-1} above the ground state greatly improves the fit to the measured temperature dependence of the excited state transitions.

0 0 0 0 4 3 0 8 0 4 9

CHAPTER IV REFERENCES

1. A. M. deGöer and B. Dreyfus, Phys. Stat. Sol. 22, 77 (1967).
2. L. J. Challis and A. M. deGöer, Phys. Letters 31A, 463 (1970).
3. M. Locatelli, P. Jaussaud, and A. M. deGöer, C. R. Acad. Sci. Paris B 270, 620 (1970).
4. R. S. Anderson, R. G. Brabin-Smith, and V. W. Rampton, J. Phys. C: Solid St. Phys. 3, 2379 (1970).
5. R. S. Anderson, C. A. Bates, and P. C. Jaussaud, J. Phys. C: Solid St. Phys. 5, 3397 (1972).
6. R. S. Anderson, C. A. Bates, P. C. Jaussaud, and V. W. Rampton, J. Phys. C: Solid St. Phys. 5, 3414 (1972).
7. R. L. Aurbach, P. L. Richards, and R. A. Forman, Bull. Am. Phys. Soc. 18, 1572 (1973) and Bull. Am. Phys. Soc. 20, 435 (1975).
8. J. H. M. Stoelinga, P. Wyder, L. J. Challis, and A. M. deGöer, J. Phys. C: Solid St. Phys. 6, L486 (1973).
9. These results have also been reported in R. L. Aurbach and P. L. Richards, Phys. Rev. (to be published).
10. Particularly useful references for group theory in this context are L. M. Falicov, Group Theory and Its Physical Applications, (University of Chicago Press, Chicago, 1966); C. J. Ballhausen, Introduction to Ligand Field Theory, (McGraw-Hill Book Co., New York, 1962); and M. Tinkham, Group Theory and Quantum Mechanics (McGraw-Hill Book Co., New York, 1964).

11. R. R. Joyce and P. L. Richards, Phys. Rev. 179, 375 (1969).

12. R. R. Joyce, Ph.D. Thesis, Dept. Phys., Univ. of Calif., Berkeley, (1970).

13. C. A. Bates and J. M. Dixon, J. Phys. C: Solid St. Phys. 2, 2209 (1969).

14. C. A. Bates and J. M. Dixon, J. Phys. C: Solid St. Phys. 2, 2225 (1969).

15. C. A. Bates, P. C. Jaussaud, and W. Smith, J. Phys. C: Solid St. Phys. 6, 898 (1973).

16. F. S. Ham, in Electron Paramagnetic Resonance, S. Geschwind, ed., (Plenum Press, N.Y., 1972).

17. A. Abragam and B. Bleaney, Electron Paramagnetic Resonance of Transition Metal Ions (Clarendon Press, Oxford, 1970).

18. D. S. McClure, J. Chem. Phys. 36, 2757 (1962).

19. J. C. Slater, Quantum Theory of Atomic Structure, (McGraw-Hill Book Co., New York, 1960), Vol. II, p. 336.

20. M. Rotenberg, R. Bivins, N. Metropolis, and J. K. Wooten, Jr., The 3-j and 6-j Symbols, (Technology Press, Cambridge, Mass., 1959).

21. E. U. Condon and G. H. Shortley, The Theory of Atomic Spectra, (Cambridge University Press, London, 1957), Chapter VI.

22. R. A. Forman, private communication.

0 0 0 4 3 0 8 0 5 0

V. γ-IRRADIATED RUBY

When ruby ($Al_2O_3:Cr^{3+}$) is irradiated with γ rays from a ^{60}Co source, color centers are formed which change its color from red to orange. In addition to the increased optical absorption,[1] the microwave EPR signal from the Cr^{3+} ions is reduced, often nearly in half.[2] If the crystal is used as a laser rod, an increased threshold is observed, with the rod finally ceasing to lase if the dosage is sufficiently high.[3] Because of the obvious technological importance of this material, many authors have attempted to study the radiation damage mechanism. It has been suggested[4] that an important radiation damage mechanism in ruby could be associated with the formation of Cr^{2+} and Cr^{4+} ions. One of the simplest of the proposed mechanisms, in that it does not require additional types of damage centers, is a charge-transfer reaction,

$$2Cr^{3+} + h\nu \rightarrow Cr^{2+} + Cr^{4+} \quad . \tag{66}$$

In this chapter I will discuss measurements made on two samples of γ-irradiated ruby. The far infrared is expected to be a useful tool to probe the properties of the damage because both Cr^{2+} and Cr^{4+} are expected to have level splittings at far infrared frequencies. Since Cr^{2+} is isoelectronic to Mn^{3+}, it is believed to have a similar level structure. Microwave studies of Cr^{4+} by Hoskins and Soffer[5] indicate that this ion can be described by a spin Hamiltonian with effective spin S=1.

$$\mathcal{H} = D[S_z^{\,2} - \frac{1}{3} S(S+1)] + \beta \underset{\sim}{H} \cdot \underset{\approx}{g} \cdot \underset{\sim}{S} \; , \tag{67}$$

where β is the Bohr magneton. From the temperature dependence of the strength of the $\Delta M=2$ transition they observed, Hoskins and Soffer were able to estimate that the zero field splitting D is about 7 cm^{-1}.

A. Short Ruby

The first sample I will discuss is a 2 cm long cylinder of ruby. The Cr concentration is about .5% by weight and the only trace impurities observed are Ba and Fe. In this sample, no low frequency (below 25 cm^{-1}) absorption lines were observed. At higher frequencies, as is shown by the spectrum labeled "0 hours bleaching" in Fig. 22 a very strong absorption line is observed at 146 cm^{-1} with weaker lines present at 236 cm^{-1}, 252 cm^{-1}, and 275 cm^{-1}. These lines are ascribed to an unknown impurity. They are associated in some way with radiation damage, however, because after high temperature annealing of the sample these lines are no longer seen. Since Ba is the largest trace impurity it is tempting to consider a mechanism involving this ion. The Ba^{2+} ion is, of course, diamagnetic. However one might consider some mechanism involving a radiation-induced local phonon mode near a Ba site or perhaps formation of Ba^{+} and surrounding color centers. These ideas are purely conjectural at this time, however, since the experimental evidence is too sparce for any definite conclusions.

It is known that under exposure to room lighting, irradiated ruby will optically bleach. That is, its color will gradually return to its original color. Although some of its other properties return to their original values, some, including the laser threshold do not.

0 0 0 0 4 3 0 8 0 5 1

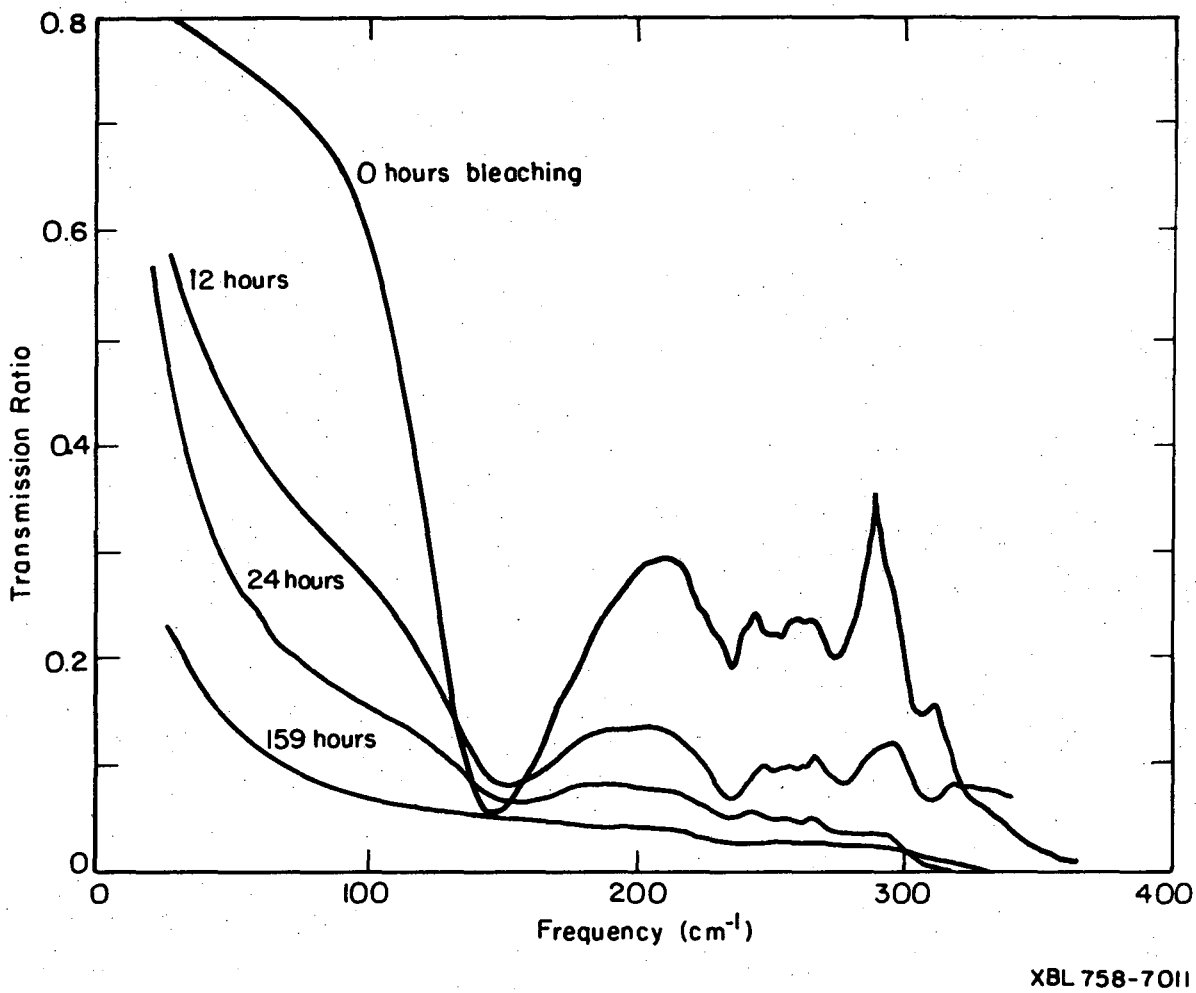


Fig. 22. Effect of Optical Bleaching on the Transmission of the Short Ruby Sample.

A high temperature annealing operation, which apparently empties deep traps, is necessary to restore the original laser threshold. We have studied the effect of optical bleaching on the infrared transmission of this sample. As is seen in Fig. 22, as the length of the optical bleaching period is increased from zero to 159 hours, the high frequency transmission falls. It appears that the strength of the unknown impurity lines is not strongly effected, but rather that additional absorption develops which masks the impurity lines. This effect is also unexplained.

B. Long Ruby

The other γ-irradiated ruby sample was a 15 cm long crystal whose c-axis made an angle of 60° to the cylinder axis. This sample was quite pure, the only spectrochemically detected impurity being Cu. The far infrared spectrum, Fig. 23(a), showed the presence of Ti^{3+}, probably at levels too small to be detected spectrochemically. In addition an absorption line was observed at 59 cm^{-1}, which disappeared upon optical bleaching. Studies of the low frequency transmission of the sample revealed weak absorption lines at 6.4 and 7.5 cm^{-1} in zero field at 4.2K. In a field of 51 kG applied along the cylinder axis, only a single line at 4.5 cm^{-1} was observed. A field on-field off ratio is shown in Fig. 23(b). The zero field lines were observed with a signal-to-noise ratio near two, so that applied field splittings would not have been observed. After several days of optical bleaching, these absorption lines were no longer observed. Since the available

0 0 0 0 4 3 0 8 0 5 2

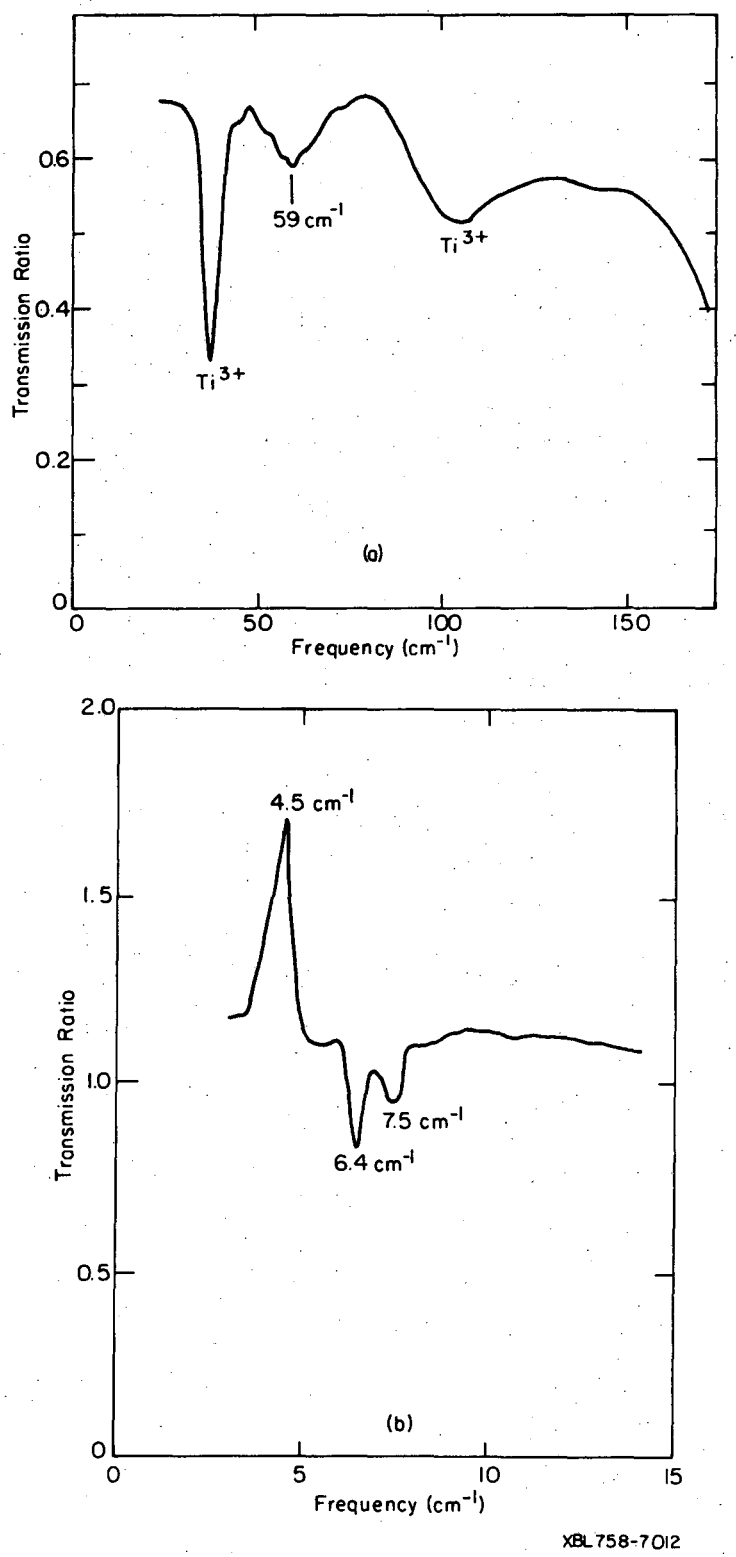


Fig. 23. Transmission of the Long Ruby Sample.

infrared data are poor, it is not possible to determine the ions responsible for the zero field lines. The 4.5 cm^{-1} line in non-zero field would seem to have a field-dependent intensity and might therefore be assigned to Cr^{2+}.

It is known that the cubic field, spin-orbit, and spin-spin interactions of Cr^{2+} in typical environments are roughly two-thirds as strong as for Mn^{3+}. If we assume that all parameters of Cr^{2+} are two-thirds their value in Mn^{3+} and further assume that the quenching factors for the two ions are the same, then the resultant level diagram is qualitatively similar to Mn^{3+}. The energy of the zero field $1 \rightarrow 5$, 6 transition is 6.5 cm^{-1}. The $1 \rightarrow 2$ transition has an energy of 4 cm^{-1} for a field of 50 kG oriented at 60° with respect to c. These results are undoubtedly fortuitous, but they suggest one possible interpretation of the data for γ-irradiated ruby. Comparison of this model with APR data[6] for Cr^{2+} yields rather poor agreement, as one would expect from such a crude estimate of model parameters.

If we assume that either the 6.5 cm^{-1} line or the 7.5 cm^{-1} line is due to Cr^{2+}, we may make a rough estimate of its concentration. If we assume that the transition probability per ion is the same for Cr^{2+} and Mn^{3+}, then by assuming that all of the Mn in sample S-18 is in the 3+ valence state and by comparing peak absorption coefficients and sample volumes we may obtain an upper limit of the Cr^{2+} concentration in our sample of γ-irradiated ruby. We find an upper limit of 5×10^{16} cm^{-3}. This represents only one part in 10^4 of the Cr in the crystal. The assumption that the transition probability per ion is the same for

0 0 0 0 4 3 0 8 0 5 3

Cr^{2+} and Mn^{3+} may not be correct. It has been found in APR measurements[7] that Mn^{3+} lines are typically ten to fourteen times as intense as Cr^{2+} lines in typical samples. If the concentrations of the two ions in typical crystals is the same, then our upper limit is at least an order of magnitude too small. The APR result could, however be interpreted as showing that the concentration of Mn^{3+} is typically higher than Cr^{2+} in Al_2O_3.

This estimate is much lower than usually quoted. Most estimates of Cr^{2+} concentration are made by assuming that the change in the Cr^{3+} EPR signal upon irradiation is directly proportional to the amount of Cr^{2+} formed.[8] The above estimate, however, suggests that the decrease in the EPR signal may be due to some other aspect of the damage mechanism and that Cr^{2+} formation is unlikely to be associated with the most important damage mechanisms in ruby.

We have studied several samples of Mg-doped ruby in the hope that Cr^{4+} formation would be favored through a change compensation mechanism, but have been unable to observe any absorption lines.

CHAPTER V REFERENCES

1. T. Marayama, Y. Matsuda, H. Kon, and H. Yonemitsu, J. Phys. Soc. Japan 18, Supplement II, 315 (1963); L. F. Vereshchagin, S. V. Starodubtsev, and M. S. Yunusov, Sov. Phys. Doklady 9, 983 (1965); G. E. Arkhangelskii, Z. L. Morgenshtern, and V. B. Neustruev, Phys. Stat. Sol. 29, 831 (1968).
2. G. E. Arkhangelskii, Z. L. Morgenshtern, and V. B. Neustruev, Phys. Stat. Sol. 22, 289 (1967) and Phys. Stat. Sol. 36, 451 (1969); R. F. Blunt, T. Chang, and M. I. Cohen, NBS Technical Note 565, 54 (1970).
3. V. R. Johnson and R. W. Grow, IEEE J. Quantum Electronics, QE3, 1 (1967); Z. L. Morgenshtern and V. B. Neustruev, Bull. Acad. Sci. USSR Phys. Ser. 32, 7 (1968); C. M. Stickley, H. Miller, E. E. Hoell, C. C. Gallagher, and R. A. Bradbury, J. Appl. Phys. 40, 1792 (1969).
4. For example, D. R. Mason and J. S. Thorp, Proc. Phys. Soc. 87, 49 (1966) and T. Maruyama and Y. Matsuda, J. Phys. Soc. Japan 19, 1096 (1964).
5. R. H. Hoskins and B. H. Soffer, Phys. Rev. 133, A490 (1964).
6. R. S. Anderson, C. A. Bates, and P. C. Jaussaud, J. Phys. C: Solid St. Phys. 5, 3397 (1972).
7. C. A. Bates, private communication.
8. For example, G. E. Arkhangelskii, Z. L. Morgenshtern, and V. B. Neustruev, Phys. Stat. Solidi 22, 289 (1967), T. Maruyama, Y. Matsuda, H. Kon, and H. Yonemitsu, J. Phys. Soc. Japan 18, 315 (1963), and A. F. Forestieri and H. H. Grimes, NASA Technical Note TN D-3379 (April 1966).

0 0 0 0 4 3 0 8 0 5 4

VI. ELECTRON-HOLE DROPS IN Ge

In the eight years since the initial observation by Rogachev and coworkers[1] of anomalous photoconductivity in Ge due to the formation of electron-hole drops (EHD) many workers have studied various theoretical and experimental aspects of drop phenomena. Pokrovskii has written an excellent review of early work in this field.[2] In addition to the free-exciton luminescence line at 714 meV, Pokrovskii and Svistunova[3] observed a new luminescence line at 709 meV at low temperatures and high levels of optical pumping. This line is due to the formation of electron-hole drops. Other studies of the luminescence[4-10] have included the existence of thresholds for droplet formation, detailed investigation of the luminescence lineshape in order to determine the condensation energy and density of carriers in droplets, a determination of the liquid-gas phase diagram, and careful studies of the optical hysteresis near threshold, which have resulted in theories for the nucleation of drops from a supersaturated exciton gas. Studies of the decay kinetics of the drops[11,12] determine a droplet lifetime of 40 μs and reveal evidence for boiling of excitons at the drop surface. Rayleigh-Gans scattering[13,14] has been used to measure drop radii of from 2 to 10 μ depending on experimental conditions. Vavilov and coworkers[15,16,17] observed far infrared emission and absorption from the EHD plasma. This work has recently been confirmed and extended to include effects due to impurities in doped Ge by Timusk and Silin.[18] A number of workers have also studied

then in the small absorption limit, this ratio is just $\alpha(\nu)d$, where d is some appropriate thickness. In Fig. 24 we have plotted a typical result. The solid line is our measurement. The other lines are the results of measurements by Timusk and Silin,[18] Murzin et al.,[20] and Kononenko, et al.,[17] normalized so that the peak heights are all the same. The vertical bars indicate the signal-to-noise ratio in various spectral regions. As can be seen the results are in agreement with earlier work. In order to interpret our result, we must digress to briefly discuss some important results of Mie theory.

B. Mie Theory

In order to understand the results of the α-drop measurements described above and the γ-drop experiments to be described below, we must first ask how an electromagnetic wave interacts with a spherical particle. In general, a particle will remove part of the incident radiation from the beam, partly through scattering and partly through absorption. We must study, therefore, the scattering and absorption due to a particle of given size and dielectric constant as a function of the frequency of the applied radiation. Then we must try to determine what is observed by the measurement. That is, we know that the absorbed radiation will not be detected. However, some fraction of the scattered radiation will be detected and some fraction of it will not, depending on the geometry of the experiment. We will first consider the derivation of scattering and absorption cross-sections.

0 0 0 0 4 3 0 8 0 5 5

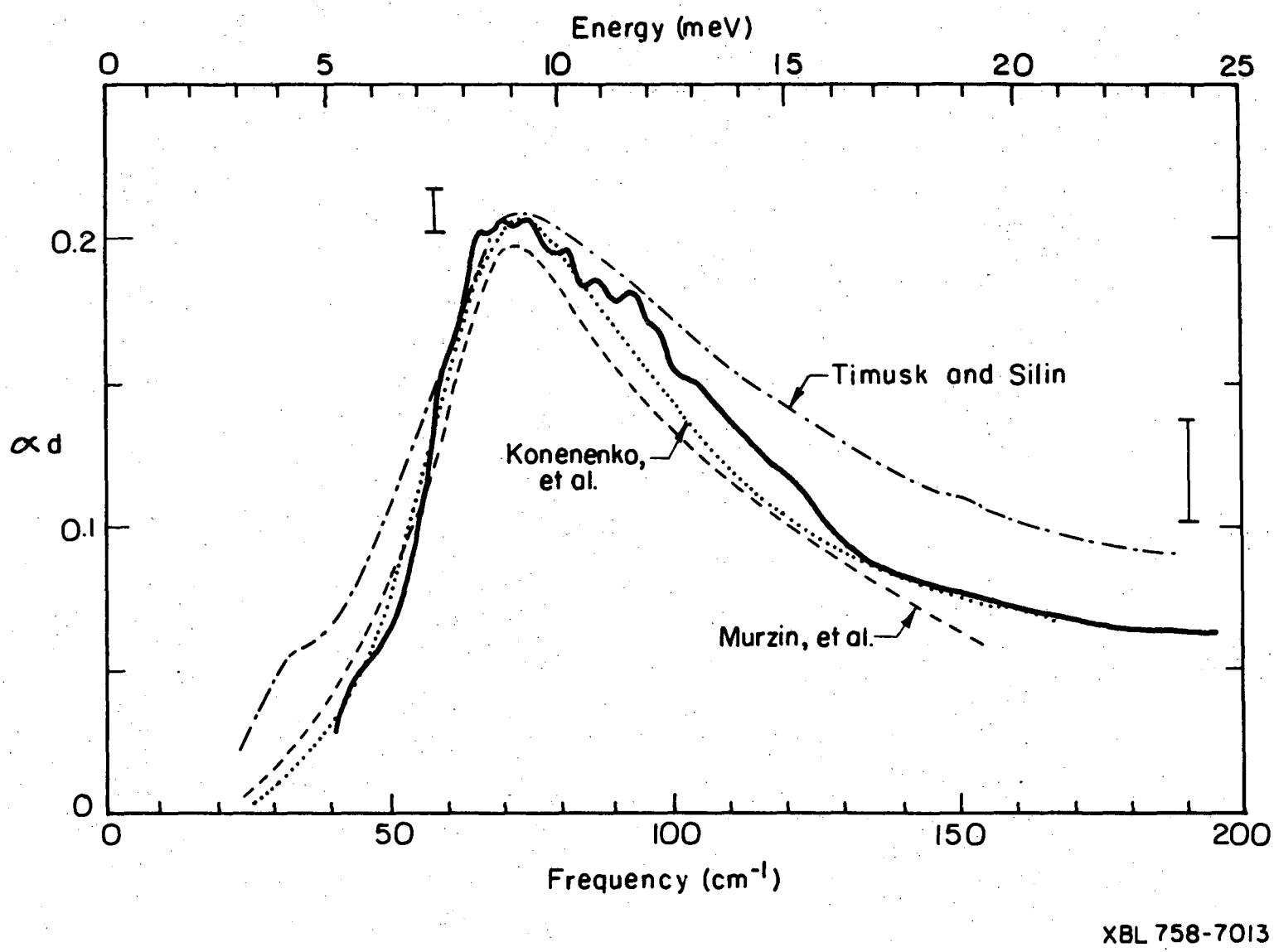


Fig. 24. Plasma Absorption of α-Drops--Comparison to other workers.

The theory of the interaction of a plane wave with a sphere of arbitrary size and dielectric constant was first discussed by Mie.[34] Modern treatments are given by Born and Wolf,[35] Jackson,[36] and van de Hulst.[37,38] We shall follow the notation of van de Hulst. Basically, the solution involves expanding the incident wave, scattered wave, and the wave inside the sphere in terms of a multipole expansion about the center of the sphere. The expansion coefficients are determined by solution of the resulting boundary value problem on the surface of the sphere. Since the full details of the theory are quite complex we will not present derivations here. Instead we will summarize some of the more important results. In these results, we will assume that the sphere is nonmagnetic ($\mu=1$), define the scattering angle θ as the angle from the proceeding beam rather than from the incident beam, and use dimensionless units throughout. This means that distances will be measured in units of $\lambda/2\pi$, where λ is the wavelength outside the sphere, and that fluxes will be measured in units of the flux geometrically incident on the sphere.

If we consider unpolarized radiation incident on the sphere, then complete information on the interaction is contained in the quantities E and $I(\theta)$. E is the extinction cross-section--that is, the fraction of flux lost by the incident wave. $I(\theta) = I_1(\theta) + I_2(\theta)$ is the flux scattered into a unit solid angle in the direction θ for radiation polarized perpendicular (I_1) and parallel (I_2) to the plane containing the incident and scattered waves. The total scattered flux is given by

0 0 0 4 3 0 8 0 5 6

$$S = \int_0^{\pi} 2\pi I(\theta) \sin\theta d\theta \tag{68}$$

and the total absorbed flux is simply

$$A = E-S \ . \tag{69}$$

If the sphere has radius r, then we may define the dimensionless parameter x by

$$x = 2\pi r/\lambda \ . \tag{70}$$

The sphere is characterized by a dielectric constant ε relative to the external medium (assumed to be vacuum). Since in our case, the sphere is imbedded in Ge, with static $\varepsilon_o = 16$, ε will be measured relative to ε_o. Also since the wavelength of the incident radiation in Ge is reduced relative to free space, we write

$$x = 2\pi\sqrt{\varepsilon_o}\ r \quad , \tag{71}$$

where ν is the wavenumber. We define the complex index of refraction in the sphere as $m \equiv \sqrt{\varepsilon}$ and introduce an auxiliary parameter y=mx.

Solving the multipole boundary value problem, we find that the coefficient of the scattered electric 2^n-pole component is

$$a_n = \frac{x\psi_n'(y)\psi_n(x) - y\psi_n'(x)\psi_n(y)}{x\psi_n'(y)\zeta_n(x) - y\zeta_n'(x)\psi_n(y)} \quad , \tag{72}$$

while the coefficient of the scattered magnetic 2^n-pole component is

$$b_n = \frac{y\psi_n'(y)\psi_n(x) - x\psi_n'(x)\psi_n(y)}{y\psi_n'(y)\zeta_n(x) - x\zeta_n'(x)\psi_n(y)} \quad . \tag{73}$$

magnetoplasma effects.[19-22] Theoretical calculations have been made of the binding energy and carrier density in drops.[23-26] Very recently a calculation of the phase diagram has been reported.[47]

It has been observed [27-30,48] that a large long-lived electron-hole drop can be formed in a sample with a large inhomogeneous stress in the ⟨110⟩ or ⟨111⟩ directions. These drops, which have diameters of from 100 to 1000 μ and luminescence decay times of 400 to 500 μsec, form at the point of maximum stress in the sample. Microwave dimensional resonances have been observed which agree well with theoretical calculations.[31] Decay kinetics and time-resolved luminescence profile studies have been performed. In addition, photographs of the drops have been taken by imaging the drop luminescence onto the surface of an infrared-sensitive vidicon tube. In order to distinguish between small drops in unstrained Ge and these large drops, the small drops are called α-drops, while the large ones are called γ-drops.

In this chapter we shall discuss results of far infrared spectroscopy of both α and γ drops.

A. Alpha-Drop Measurements

Measurements of α-drop samples in the region from 30 to 200 cm^{-1} were made using an optical cold-finger cryostat which made possible simultaneous measurements of far infrared transmission and near infrared luminescence. The cold-finger dewar was designed with a pump-light port in the bottom of the dewar, a near infrared luminescence port on one side of the dewar, and two far infrared ports on two

0 0 0 0 4 3 0 8 0 5 7

opposite sides. The pump-light and luminescence ports were sealed with glass and quartz windows respectively. The far infrared ports were sealed with polyethylene windows. Pump light from a Hg arc lamp was focussed on the bottom face of the sample. Luminescence from a side face of the cubic sample was collected, chopped, and imaged on the entrance slit of a 1/4 m grating spectrometer. The output was detected using a Ge photodetector operated at 77K. Two sources of background radiation had to be removed in order to insure that only the luminescence signal reached the near infrared detector. The first of these was visible light which could pass through the spectrometer in second order. This was removed with a Corning 7-56 filter. The second was near infrared radiation from the pump lamp. A water filter was used to remove this radiation from the pump lamp input. Far infrared radiation was coupled into and out of the sample using Winston light concentrators.[32] Cold wedged crystal-quartz filters removed excess background radiation. Far infrared radiation was detected using a doped-Ge bolometer in an external cryostat.

The sample was a polished and etched cube of high purity Ge[33] approximately 5 mm on a side. In order to insure good thermal contact between the sample, which was mounted in the main vacuum space of the dewar, and the liquid helium vessel, the sample was soldered with In to a thick Cu heat sink which was bolted to the bottom of the helium chamber. Leads were attached to the sample so that photoconductivity could also be measured. Cooling the sample was one of

the most difficult problems which had to be overcome in the experimental design. In early stages of the experiment, photothermal conductivity was observed from Al impurities. The presence of this absorption showed that the sample temperature was about 8K. This was well above the EHD critical temperature $T_c \approx 6.5K$. By minimizing the background radiation power incident on the sample and by increasing the thermal conductance between the sample and the liquid helium pot, we were able to reduce the equilibrium sample temperature to a few tenths of a degree above the 4.2K bath temperature.

With the experimental arrangement described above several types of far infrared measurements could be made, depending on which signal was chopped. The output of the far infrared Michelson interferometer could be chopped or used dc. The pump light could also be run either ac or dc, depending on the location of a light chopper. Although feasible, double demodulation schemes were not used. Rather, selections were made of the appropriate signals to be chopped.

The presence of pump light at energies above the gap changes the sample resistance from the megohm range when dark to the kilohm range. Therefore chopping the pump light would result in a very large background signal in a photoconductivity measurement, and these experiments were done using dc pump light. For far infrared transmission measurements there are two considerations. Since Vavilov[15] observed plasma emission, we expect plasma emission to be present in our experiment. This radiation does not pass through the spectrometer and will, therefore, appear as a background signal. If we use dc

0 0 0 0 4 3 0 8 0 5 8

pump-light, the emission signal will be present at dc and will be rejected by the lock-in amplifier. If we use ac pump light, the emission will be a background signal at the chopping frequency. If the plasma emission is a significant contribution, it would be advantageous to use dc pump light. On the other hand, if we chop the Michelson output and if droplet absorption is small, we will be looking for a small absorption on a large background transmission, which requires a good signal-to-noise ratio. From this point of view we would prefer to chop the pump light because this would "chop the effect" and give a signal which is more directly related to drop properties. The choice between these two conflicting considerations must therefore be based on the magnitudes of the drop absorption and emission signals.

Several experiments were first performed at 4.2K. Photoconductivity was measured both with and without pump light. The amplitude and spectral shape of the photoconductive response were found to be functions of light level and bias conditions. Far infrared absorption was not observed, although a background signal was seen when chopping the pump light. Luminescence was not observed, either from drops or from free excitons, presumably because of improper alignment. These experiments were unsatisfactory because it is possible to explain the negative far infrared transmission results regardless of whether or not drops were formed. If they were not formed, we should not, of course, expect to see far infrared absorption. The background signal observed when the pump light was chopped could be far infrared light from the Hg arc lamp. If drops were formed, the lack of

far infrared absorption is most likely due to poor geometry in the experiment. Since the pump light penetrates only a few microns into the sample, drops would be formed only in a thin layer very near the bottom surface of the crystal. Only a small fraction of the incident far infrared radiation would pass through this layer. Therefore the efficiency for the detection of far infrared absorption is low. The photoconductivity results remain unexplained.

With the apparatus configured as above, measurements were made with the bath pumped to approximately 1.8K. Since the sample was soldered with In to the heat sink, the sample temperature was probably limited by the poor thermal conductivity of In below its superconducting transition temperature. Although we were unable to obtain a reliable measurement, we estimate that the sample temperature was about 3.5K. Under these conditions, near infrared drop luminescence was observed, confirming the existence of α-drops. Photoconductivity measurements were made which were, as before, light level and bias dependent and which have as yet escaped interpretation. No difference was observed between far infrared transmission measurements at 4.2K and at this temperature, probably because of the geometry problem discussed above.

In order to remedy the poor experimental geometry, modifications were made to the apparatus to enable the pump light to illuminate the surface of the crystal upon which the far infrared radiation was incident. A small hole was drilled in the elbow of the incident-beam light pipe so that the pump light could be focussed through this hole onto the sample. A small glass cover slip was epoxied over the

0 0 0 0 4 3 0 8 0 5 9

hole to maintain the vacuum seal. In order to cut down on losses in the incident light-pipe optics, the input light pipe and Winston cone were silver plated and polished.

Measurements of the far infrared transmission were make using this improved apparatus and absorption was observed. Using dc pump light, a five percent peak absorption was seen. Since this absorption was weak, it was more profitable to chop the pump light. A large background signal was observed with a small additional signal which was dependent on the interferometer path-difference. This small signal contributed only about 10% to the total signal observed. Using the lock-in offset adjustment, this signal was measured at a number of different light levels. The instrumental response was removed by dividing the observed spectrum by a spectrum taken without pump light using the chopped output of the interferometer.

We may interpret this ratio in the following way. Suppose that the instrumental response of the system including the transmission of the sample in the absence of pump light is $B(\nu)$. When the chopper is open we observe $B(\nu)\ T(\nu) + E$, where $T(\nu)$ is the change of transmission of the sample due to the pump light and E is the background signal (assumed to be drop emission). Since the emission does not pass through the spectrometer, its frequency dependence is not observed. The spectrum measured with ac pump light is, then, $B(\nu)[1-T(\nu)]$. The spectrum measured without pump-light is just $B(\nu)$. Therefore the ratio measures $1-T(\nu)$. If we may characterize the drop transmission by a Beer's Law absorption coefficient, $T(\nu) = e^{-\alpha(\nu)d}$,

where ψ and ζ are related to the Bessel functions of half-integral order by

$$\psi_n(x) = \left(\frac{\pi x}{2}\right)^{1/2} J_{n+\frac{1}{2}}(x) ,$$

$$\chi_n(x) = (-1)^n \left(\frac{\pi x}{2}\right)^{1/2} J_{-n-\frac{1}{2}}(x) , \tag{74}$$

$$\zeta_n(x) = \psi_n(x) + i\chi_n(x) .$$

The primes denote derivatives. The extinction cross-section is calculated by evaluating the Poynting vector. The result is

$$E = \frac{2}{x^2} \sum_{n=1}^{\infty} (2n+1) \left[\mathcal{R}e(a_n) + \mathcal{R}e(b_n)\right] , \tag{75}$$

where $\mathcal{R}e$ stands for real part. Evaluating some integrals, we may also calculate the scattering cross-section,

$$S = \frac{2}{x^2} \sum_{n=1}^{\infty} (2n+1) \left[|a_n|^2 + |b_n|^2\right] . \tag{76}$$

The absorption cross-section A is calculated using Eq. (69).

Although the expressions for a_n and b_n are in general exceedingly complicated, these are two limiting cases which will be of considerable interest. These are the limits when the drop is small compared to a wavelength ($x \ll 1$) and when it is large compared to a wavelength ($x \gg 1$). In both these cases, asymptotic expressions for the Bessel functions are useful.

0 0 0 0 4 3 0 8 0 6 0

For the case $x \ll 1$, only the first few terms are important. Van de Hulst[37] has calculated E and S in this limit including electric and magnetic dipole and electric quadrupole terms. He finds that

$$E = -\mathrm{Im}\left\{4x\left(\frac{\varepsilon-1}{\varepsilon+2}\right) + \frac{4}{15}x^3\left(\frac{\varepsilon-1}{\varepsilon+2}\right)^2 \frac{\varepsilon^2+27\varepsilon+38}{2\varepsilon^2+3}\right\}$$

$$+ x^4\,\mathcal{R}e\left\{\frac{8}{3}\left(\frac{\varepsilon-1}{\varepsilon+2}\right)^2\right\}, \tag{77}$$

and

$$S = x^4\,\frac{8}{3}\left|\frac{\varepsilon-1}{\varepsilon+2}\right|^2 . \tag{78}$$

The scattering cross-section is small. The x^4 term indicates that it represents Rayleight scattering. The absorption cross-section is

$$A = -\mathrm{Im}\left\{4x\left(\frac{\varepsilon-1}{\varepsilon+2}\right) + \frac{4}{15}x^3\left(\frac{\varepsilon-1}{\varepsilon+2}\right)^2 \frac{\varepsilon^2+27\varepsilon+38}{2\varepsilon^2+3}\right\} - \frac{16}{3}x^4\left\{\mathrm{Im}\left(\frac{\varepsilon-1}{\varepsilon+2}\right)\right\}^2 . \tag{79}$$

The leading term is the most interesting. Since the flux A is in units of the geometrically incident flux, the total absorbed flux is proportional to $V\nu\ \mathrm{Im}\left(\frac{\varepsilon-1}{\varepsilon+2}\right)$, where V is the volume of the drop. This expression was used by Vavilov[15] to characterize the droplet absorption.

In the case of spheres large compared to a wavelength, $x \gg 1$, we may make use of the well-known asymptotic expressions for the spherical Bessel functions

$$\psi_n(x) = \sin(x - \frac{n\pi}{2}), \tag{80}$$

$$\chi_n(x) = \cos(x - \frac{n\pi}{2}) .$$

Using these expressions, we may express a_n and b_n by

$$a_n = \frac{1}{1 + i\xi_n^a} \quad , \tag{81}$$

and

$$b_n = \frac{1}{1 - i\xi_n^b} \quad ,$$

where

$$\xi_n^a = \frac{(m+1)\cos(m-1)x - (m-1)(-1)^n \cos(m+1)x}{(m-1)\sin(m-1)x - (m+1)(-1)^n \sin(m+1)x} \quad ,$$

and

$$\xi_n^b = \frac{(m+1)\cos(m-1)x + (m-1)(-1)^n \cos(m+1)x}{(m-1)\sin(m-1)x + (m+1)(-1)^n \sin(m+1)x} \quad . \tag{82}$$

When m is complex, the trigonometric functions become combinations of trigonometric and hyperbolic functions. Using Eqs. (77) and (78) we write

$$E = \frac{2}{x^2} \sum_{n=1}^{\infty} (2n+1) \left[\frac{1-\mathrm{Im}\ \xi_n^a}{1-2\mathrm{Im}\xi_n^a + |\xi_n^a|^2} + \frac{1+\mathrm{Im}\ \xi_n^b}{1+2\mathrm{Im}\xi_n^b + |\xi_n^b|^2} \right] , \tag{83}$$

and

$$S = \frac{2}{x^2} \sum_{n=1}^{\infty} (2n+1) \left[\left(1-2\mathrm{Im}\xi_n^a + |\xi_n^a|^2\right)^{-1} + \left(1+2\mathrm{Im}\xi_n^b + |\xi_n^b|^2\right)^{-1} \right] . \tag{84}$$

Using Eq. (69) we calculated the absorption cross-section to be

$$A = \frac{2}{x^2} \sum_{n=1}^{\infty} (2n+1) \left[\frac{-\mathrm{Im}\ \xi_n^a}{1-2\mathrm{Im}\xi_n^2 + |\xi_n^a|^2} + \frac{\mathrm{Im}\ \xi_n^b}{1+2\mathrm{Im}\xi_n^b + |\xi_n^b|^2} \right] . \tag{85}$$

0 0 0 0 4 3 0 8 0 6 1

There are several comments that should be made here. First, by analogy with quantum mechanical scattering problems, it can be shown that the a_n and b_n fall rapidly to zero when $n + \frac{1}{2} > x$. Therefore the infinite sums may be terminated at n = integer part of x. Since the expressions for the a_n and b_n are exact within the approximation of Eq. (80), it is clear that the fact that a_n and b_n do not fall to zero near n=x is due to the approximation of Eq. (80), which fails when x approaches n. Debye[46] has developed considerably more complicated expressions for ψ_n and χ_n which would improve the accuracy of Eqs. (83) - (85) at the expense of a considerable increase in complexity of the computer calculation. We have not yet attempted to implement his expressions. Hence we expect that the expressions in Eqs. (83) - (85) are only approximate. Second, it is clear that in the large sphere limit that there is significant scattering. In the small sphere limit the absorption dominated the scattering because the leading term in A was proportional to x while the leading term in S was proportional to x^4. Here we expect S and A to be comparable in magnitude. As we shall see, the role of scattering complicates the interpretation of results for large electron-hole drops.

C. Interpretation of Alpha-Drop Results

We now wish to use the results of Mie theory to calculate the interaction of far infrared radiation with an α-drop. We first observe that within the framework of Mie theory, an α-drop is small. If we consider a drop with a radius of 2μ, then x ranges from .10 at 20 cm^{-1} to 1.0 at 200 cm^{-1}. The absorption should be given correctly by Eq. (79) for x values as large as 2 due to the inclusion of the magnetic dipole and electric quadrupole terms. If we consider only the leading term of Eq. (79), which is the electric dipole contribution, inaccuracies should enter near x=1.

In order to calculate A, we need to know the dielectric constant appropriate to α-drops. Since the drops are electron-hole plasmas, we use the simple dielectric function of a damped plasma,

$$\varepsilon = 1 - \frac{\omega_p^{\ 2}}{\varepsilon_o \omega(\omega + i/\tau)} \quad , \tag{86}$$

where we have written the dielectric constant relative to the static dielectric constant ε_o of Ge. The plasma frequency,

$$\omega_p^{\ 2} = \frac{4\pi n e^2}{m^*} \quad , \tag{87}$$

incorporates a reduced mass which takes into account the electrons, light holes, and heavy holes. We have characterized the plasma damping by a relaxation time τ. We can obtain a feeling for the effect of this dielectric constant by substituting it into the electric dipole component of A,

0 0 1 0 4 3 0 8 0 6 2

$$A_{el} = -4x \text{ Im} \left(\frac{\varepsilon-1}{\varepsilon+2}\right) \quad . \tag{88}$$

Writing

$$\omega_o^{\ 2} = \frac{\omega_p^{\ 2}}{3\varepsilon_o} \tag{89}$$

we have

$$A_{el} = \frac{\omega^2 \ \omega_o^{\ 2}/\tau}{(\omega^2-\omega_o^{\ 2})^2 + (\frac{\omega}{\tau})^2} \quad . \tag{90a}$$

In wavenumber units,

$$A_{el} = \frac{\nu^2 \ \nu_o^{\ 2}/\tau}{(\nu^2-\nu_o^{\ 2})^2 + \left(\frac{\nu}{2\pi c\tau}\right)^2} \quad , \tag{90b}$$

where $\nu_o = 2\pi c\omega_o$. Notice that Mie theory has made a marked change in the spectral shape. Rather than the usual Drude edge, we see a resonance at $\omega_o = \omega_p/\sqrt{3\varepsilon_o}$. This is due to the $\varepsilon+2$ resonant denominator in the expression for A. This denominator is also responsible for the shift in peak frequency from $\omega_p/\sqrt{\varepsilon_o}$, where the dielectric function crosses zero, to $\omega_p/\sqrt{3\varepsilon_o}$. Using a plasma density of 2.3×10^{17} cm^{-3} and the appropriate reduced mass m* = .078 m, we obtain ν_o = 74.2 cm^{-1}, which is in excellent agreement with the observed value of 74.5 cm^{-1}. The full expression for A (Eq. (79)) has been calculated using a relaxation time of 1×10^{-13} sec and is shown as the dotted curve in Fig. 25. The very short value of τ was chosen to increase the half-width of the resonance so as to give a better fit to the high frequency tail of the data. The solid line is our measurement. As can be seen, the high frequency tail of the absorption is in poor agreement with the

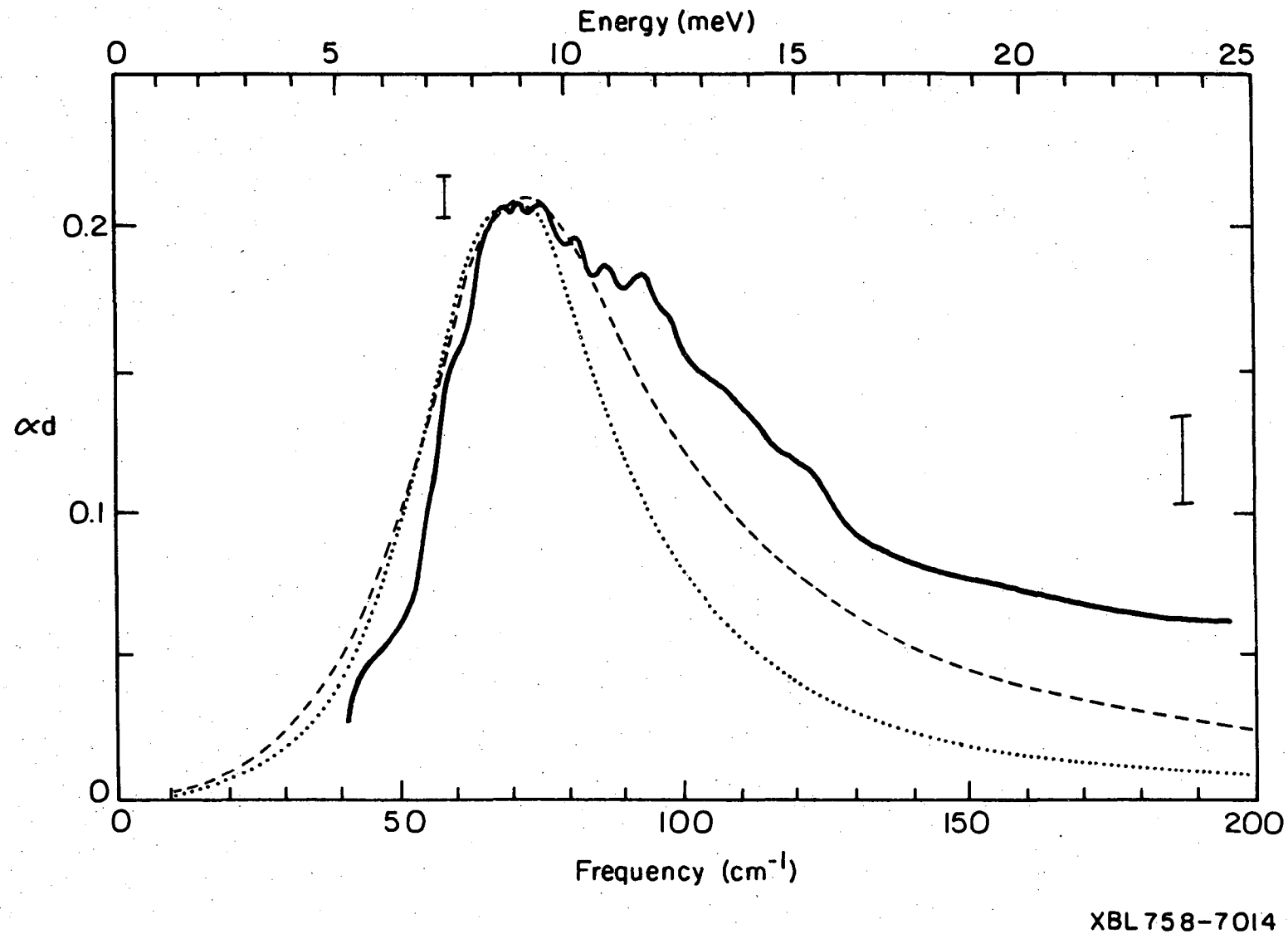


Fig. 25. Plasma Absorption of α-Drops--Comparison with Theory.

0 0 0 0 4 3 0 8 0 6 3

with the calculation.

The fit is poor because we have neglected a very important contribution to the dielectric function. As was first pointed out by Rice[39] and by Rose and Shore,[40] interband transitions play a major role in the far infrared absorption. Formation of electron-hole drops results a partial depopulation of the valence band inside the drop. Because of this depopulation, vertical optical transitions are allowed between heavy and light hole bands. These transitions give rise to extra contributions to the dielectric constant. Following a calculation by Combescot and Nozieres[41] we write the interband contribution to ε as

$$\varepsilon_{inter} = \frac{4\pi e^2}{\varepsilon_o}\left(\frac{1}{m_2} - \frac{1}{m_1}\right)\sum_k \frac{f(E_{k_1}) - f(E_{k_2})}{(E_{k_1} - E_{k_2})^2 - \omega(\omega + i/\tau)} \tag{91}$$

where f is the fermi function and the heavy and light hole bands are characterized by energies E_{k_1}, E_{k_2} and masses m_1, m_2. Replacing the sum by an integral, we write

$$\varepsilon_{inter} = \left(\frac{e}{\pi}\right)^2 I \ , \tag{92a}$$

where

$$I = \int d\Omega \left[\frac{1}{2m_2} - \frac{1}{2m_1}\right] \int_{k_{F_1}}^{k_{F_2}} \frac{k^2 dk}{(E_{k1} - E_{k2})^2 - \omega(\omega + i/\tau)} \ . \tag{92b}$$

The limits of integration are the fermi wavevectors for the light and heavy hole bands. In Ge the valence bands are warped. This has several

effects. First the masses are no longer scalers, which complicates the integration over angle. Secondly, the angular dependence of the masses makes the energy eigenvalues complicated functions of angle. Finally the energy band warping means that the fermi level will intersect the valence bands at different values of k in different directions. Therefore the limits of integration of the integral over wavevector have angular dependence as well. The full solution of the problem is complicated and must be performed numerically. If we assume that the bands are spherical, however, the solution is analytic.

In the spherical band approximation, we perform the angular integration trivially. Defining

$$\frac{1}{m} = \frac{1}{m_2} - \frac{1}{m_1} , \tag{93}$$

we write

$$I = 4m\pi \int_{k_1}^{k_2} \frac{k^2 dk}{k^4 - a^4} , \tag{94}$$

where $a^4 \equiv 4m^2\omega(\omega + i/\tau)$. The limits of integration are determined by requiring that the fermi level is the same for the two bands and that the total number of states removed from the bands is determined by the density of the drop. Satisfying these two conditions we get

$$k_1 = \left[\frac{3\pi^2 n}{1 + \left(\frac{m_2}{m_1}\right)^{3/2}} \right]^{1/3} , \tag{95a}$$

0 0 0 4 3 0 8 0 6 4

$$k_2 = \sqrt{\frac{m_2}{m_1}}\ k_1 \ . \tag{95b}$$

The integral is evaluated by decomposing it into a sum of partial fractions.

$$I = \frac{m\pi}{a} \int_{k_1}^{k_2} \left[\frac{1}{k-a} - \frac{1}{k+a} + \frac{i}{k+ia} - \frac{i}{k-ia} \right] dk \ . \tag{96}$$

We write $a = \mu+i\nu$ and $d=\mu^2+\nu^2$ and separate the integrals into real and imaginary parts. Performing the integrals we get

$$I = \frac{m\pi}{d} (\mu-i\nu) \left\{ \frac{1}{2} \ln \left(\frac{k^2-2k\mu+d}{k^2+2k\mu+d} \right) + \frac{i}{2} \ln \left(\frac{k^2-2k\nu+d}{k^2+2k\nu+d} \right) \right.$$

$$\left. + \tan^{-1} \left(\frac{2k\mu}{k^2-\mu^2-\nu^2} \right) - i \tan^{-1} \left(\frac{2\mu\nu}{k^2-\mu^2+\nu^2} \right) \right\} \Bigg|_{k_1}^{k_2} \ . \tag{97}$$

This expression and the corresponding expression for warped bands have been evaluated numerically. The dashed line in Fig. 25 is the result of a warped band calculation using a density of 2.3×10^{17} cm^{-3} and a lifetime of 1×10^{-13} sec. Although there are still discrepancies between theory and experiment, inclusion of the interband terms improves the fit. Rice[39] has performed an improved calculation which is in excellent agreement with experiment. In that calculation he has included the angular dependence of the optical transition matrix elements. The calculation described here assumed that these matrix elements were spherically symmetric.

D. Gamma Drop Experiments

For several reasons our experiments on γ-drops could not be carried out using the cold-finger cryostat. Firstly the application of the proper uniaxial stress was quite difficult in the cold-finger cryostat. A permenent stress sample holder has been devised by J. P. Wolfe which consists of a small Kel-F ring. The sample is held in the ring by a nylon screw. Stress is applied at room temperature. Cooling the sample, which results in differential contraction between sample and holder, increases the amount of stress. With this type of holder, the sample may be cycled between room temperature and low temperature indefinitely without adverse effect on the formation of a γ-drop. Clearly it is highly desirable to use a sample mounted in such a holder. Unfortunately the cold-finger dewar is not suited to its use. Since metal contacts inhibit γ-drop formation, presumably because of the creation of Schottky barriers, the sample cannot be soldered to a heat sink. Therefore cooling the sample becomes a major problem. In addition magnetoplasma experiments cannot be performed in the cold-finger cryostat.

For the above reasons, a special insert was constructed for the cryostat described in Chapter IV. The usual 1.1 cm light pipe which normally carries radiation from the top of the dewer to the sample could not be used because of excessive losses in the visible pump beam and particularly the near infrared luminescence. The isotropic luminescence radiation is collected by a cone. The radiation leaving

0 0 0 0 4 3 0 8 0 6 5

the cone has an f/1.5 angular distribution for 1.1 cm optics. This beam would strike the walls of the light pipe about 15 times per meter of travel. Considering the absorptivity of metals in the near infrared, the resultant loss is unacceptably high. In order to cut down on this loss, a $\frac{7}{8}$ - inch stainless steel light pipe was chosen. Since the radiation in this pipe is closer to f/4, there would only be about 6 bounces in the meter length of pipe. Unfortunately stainless steel has a rather poor infrared reflectivity so a roll of 5-mil aluminized Mylar was inserted in the tube to act as the reflecting surface for the radiation. A cone adapter converted the output of the interferometer to a beam of the required diameter. A silver-plated two piece cone concentrated radiation onto the sample. It consisted of a straight cone which adapted between $\frac{7}{8}$ - and $\frac{1}{2}$ - inch diameters and a Winston cone to concentrate radiation on the sample. Pump light from a 150W incandescent projection lamp was applied to the sample through a quartz window in the side of the light pipe. The pump light was reflected from a small diagonal mirror in the insert down onto the sample.

The sample chosen for these measurements was strained along the ⟨110⟩ direction and was shown photographically[29] to contain a γ-drop. The first measurements were made at 4.2K with the sample immersed in helium. Chopping the MI, spectra were measured with and without dc pump light. A typical transmission ratio (light-to-dark) is shown in Fig. 26(a). The transmission is seen to decrease with increasing wavenumber. We believe that this result is due to sample heating. Estimates of the sample temperature rise due to the absorption of pump

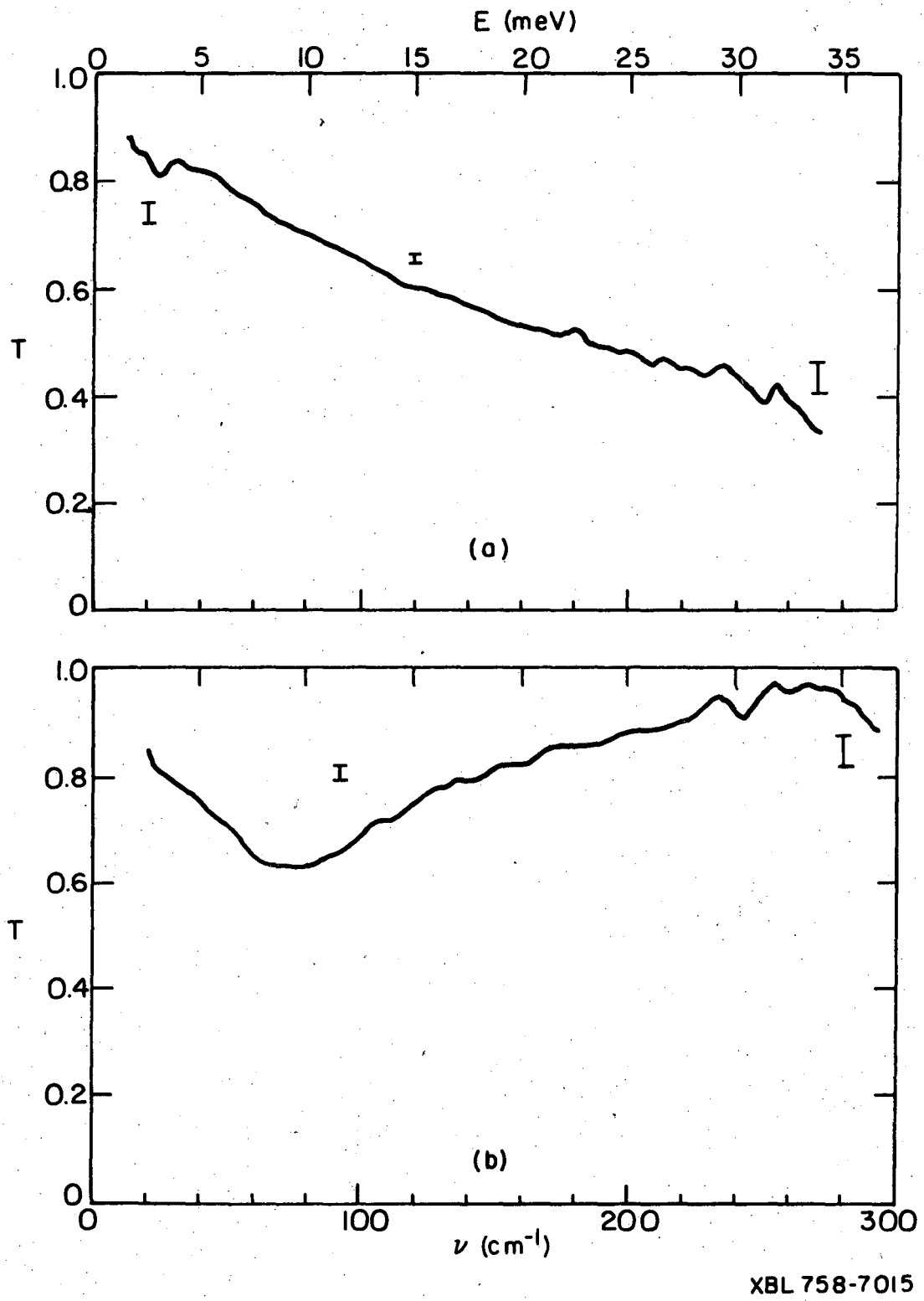


Fig. 26. Transmission of γ-Drop Sample at 4.2K and 1.8K.

0 0 0 4 3 0 3 0 6 6

light indicate that the sample temperature was probably 4.5 to 5.0K. Unpublished calculations by Vashista[42] show that the application of stress reduces the binding energy and hence the critical temperature of EHDs. His calculations indicate that the critical temperature of the γ-drop in our experiment may have been as low as 4.2K. The sample might have been above the critical temperature.

An alternate explanation of our result is that we may have used too much pump power. In γ-drop imaging experiments performed at 1.8K, it has been observed[49] that excessive pump power will destroy a γ-drop. This effect is not understood, because the threshold pump powers are too low to heat the sample above the critical temperature. The temperature dependence of this threshold power has not been studied. It is reasonable to suppose, however, that as the sample temperature approaches the critical temperature, the threshold power decreases. If this is the case, our power levels may have been above threshold.

Under the same pump light conditions at temperatures below the lambda point an entirely different transmission spectrum is observed, as is shown in Fig. 26(b). This transmission may be characterized by a frequency-dependent extinction. Unlike the α-drop, where the observed extinction was dominated by absorption, the γ-drop extinction contains a major scattering contribution. In our experiment not all of the scattering will contribute to the observed extinction because some fraction of the scattered light will be collected by the output optics and observed. The angular-distribution of the far infrared radiation in the disc-shaped sample (including multiple reflections from the

surface of the disc) makes the calculation of this fraction complicated. The following simple argument, however, suggests that most of the scattered light is collected.

The drop is fairly small compared to the sample as a whole. The drop radius is estimated to be about 100μ from photographs under similar pumping conditions. The sample is a 4 mm diameter disc, 1.35 mm thick. Hence the drop occupies only 2.5×10^{-4} of the sample volume. Furthermore the sample forms a cavity in which scattered light might be expected to be reflected several times. We intuitively expect, therefore, that much of the scattered light will eventually be collected by the output optics. Therefore the scattering should not contribute strongly to the observed extinction. This should also be true for α-drops, where the scattering is already small. We expect that the observed extinction should be well characterized by the absorption. We contrast this with the case of a drop which occupies a substantial fraction of the sample. In this case the major component to the sample transmission would be forward scattering. The observed extinction would contain contributions from both absorption and large angle scattering.

If it is reasonable to assume that the observed extinction is dominated by absorption, then it is appropriate to represent the data in terms of an absorption coefficient. We therefore calculate the negative of the natural logarithm of the light-dark ratio. This quantity is proportional to αd in the limit that the data may be characterized by an absorption coefficient alone. If another model is found for the

0 0 0 4 3 0 8 0 6 7

observed extinction, then the calculated quantity is not αd. However it is still a well-defined quantity in terms of which the data may be discussed.

A typical result for the γ-drop is shown as the solid line in Fig. 27. For comparison, the dotted line is the spectrum of an α-drop. The two spectra are surprisingly similar, the major difference being the extra γ-drop absorption at low frequency. In order to understand the similarities and differences between the observed α and γ-drop absorption, we must do two things. First we must show that the data can be interpreted in terms of the expected properties of a γ-drop. Secondly we must review the physics of γ-drops in order to show that alternate explanations for the observed absorption are inconsistent with these and previous experiments.

E. Gamma-Drop Interpretation

We will assume that the γ-drop has a radius of 100μ. Photographs (Fig. 28) of the drop using similar pump light levels as in the far infrared measurements show that this estimate is realistic. Absorption has been observed in the range from 5 to 300 cm^{-1} so that Mie theory parameter x will range from 1.3 to 75. For most of the region of interest the large sphere approximation of Mie theory will be appropriate. Unfortunately the Mie calculations are quite laborious in the large sphere limit. A qualitative and a fairly good quantitative understanding of the absorption of a large sphere can be achieved, however, using a simple model.

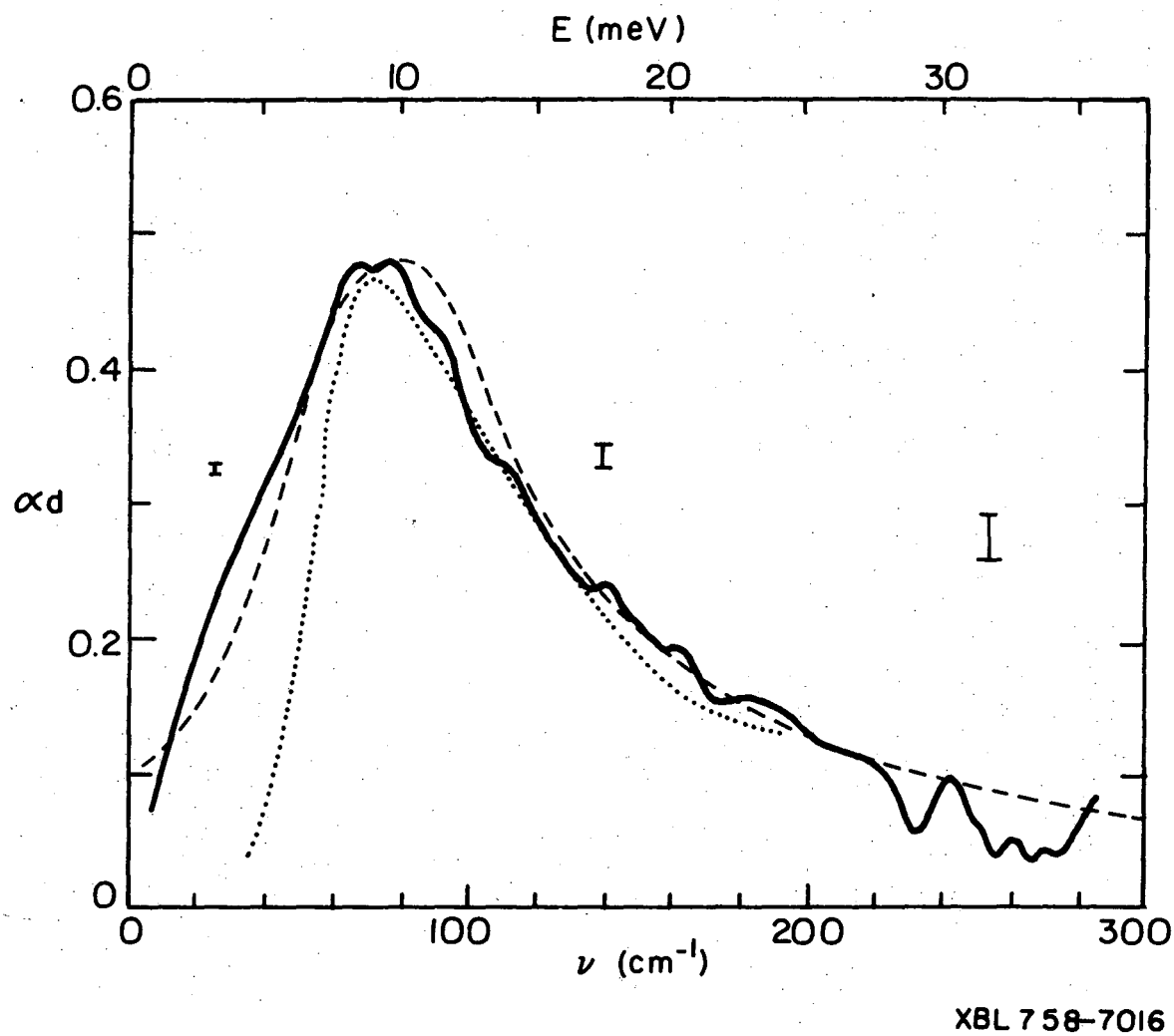
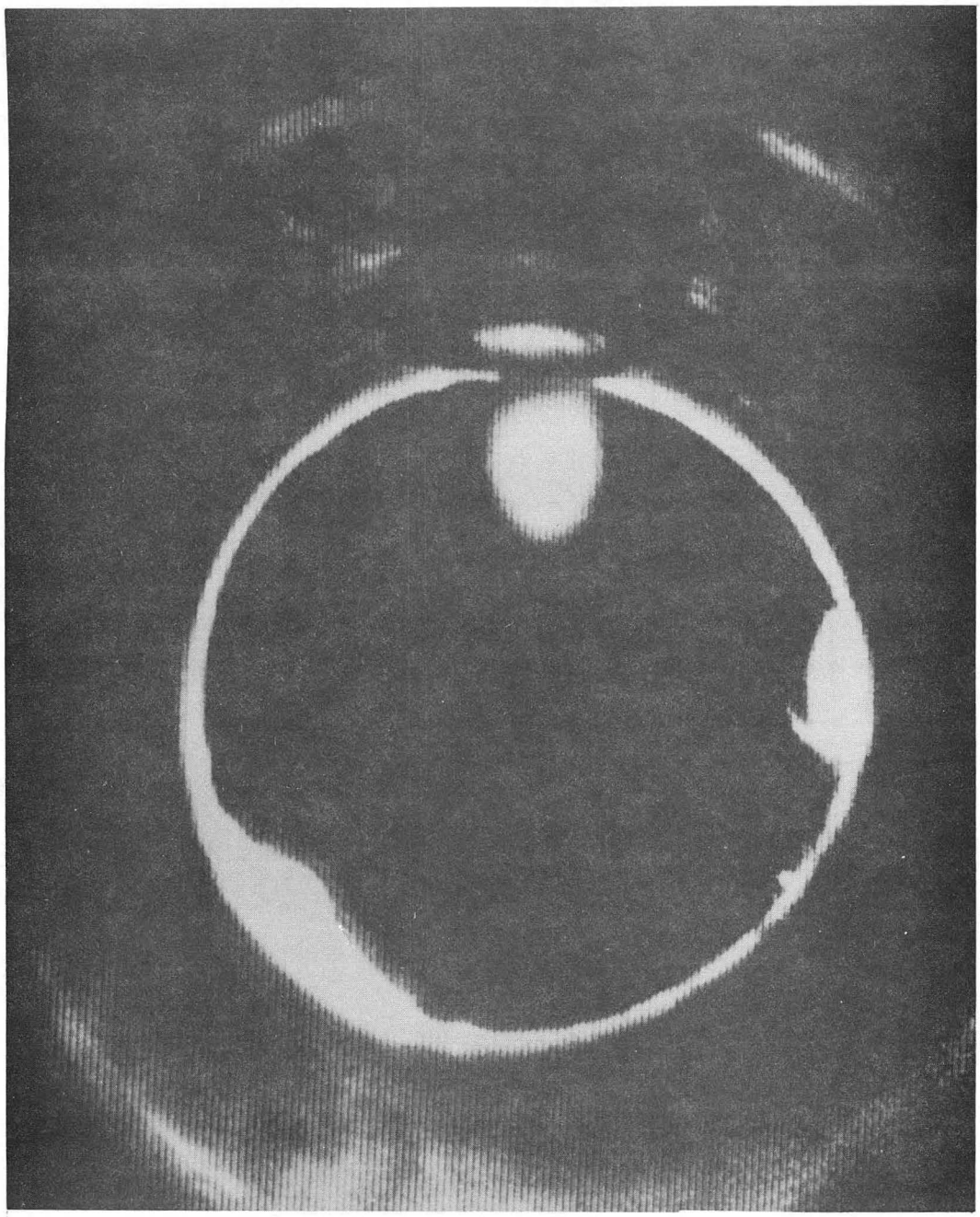


Fig. 27. Plasma Absorption of γ-Drop Sample--Comparison with Theory.

0 0 0 4 3 0 8 0 6 8



XBB 758-5810

Fig. 28. Photograph of the γ-Drop Sample.

If a plane wave of unit intensity is normally incident on a bulk material characterized by a dielectric constant ε relative to the surrounding medium (dielectric constant ε_o), the fraction absorbed is just

$$(1-R)\ [1 - \exp(-\alpha d)], \tag{98}$$

where $R = \left|\frac{n-1}{n+1}\right|^2$, $n=\sqrt{\varepsilon}$, and $\alpha = \frac{2\omega\sqrt{\varepsilon_o}}{c}$ Im n. We write $\varepsilon = \varepsilon_1 + i\varepsilon_2$. The total absorption due to the sample is

$$A = d^2(1-R)[1 - \exp(-\alpha d)]\ , \tag{99}$$

including the finite size of the sample. The absorption calculated using this simple model agrees rather well with Mie theory calculations in the large sphere limit.

If we ignore the interband term for the moment we may calculate the absorption due to the plasma dielectric function, Eq. (86). We find that

$$1-R = \frac{2\sqrt{\sqrt{\varepsilon_1^2 + \varepsilon_2^2} + \varepsilon_1}}{\sqrt{\varepsilon_1^2 + \varepsilon_2^2} + 1 + \sqrt{\sqrt{\varepsilon_1^2 + \varepsilon_2^2} + \varepsilon_1}}\ , \tag{100}$$

where

$$\varepsilon_1 = 1 - \frac{\omega_p^2\tau^2}{\varepsilon_o(\omega^2\tau^2+1)}\ , \qquad \varepsilon_2 = \frac{\omega_p^2\tau}{\varepsilon_o\omega(\omega^2\tau^2 + 1)}\ . \tag{101}$$

In general these expressions are exceedingly complicated and must be evaluated numerically. We find, however, that this expression has a well-defined maximum near $\omega_p^* = \omega_p/\sqrt{\varepsilon_o}$ which we may easily understand.

0 0 0 0 4 3 0 8 0 6 9

Near ω_p^* the reflectivity is changing from the usual metallic region below ω_p^* to the ultraviolet transparency region above ω_p^*. Therefore 1-R is rising near ω_p^*. The absorption coefficient is falling in this region since it is proportional to $1/\omega^2$ for small ω. Therefore the $1-\exp(-\alpha d)$ term is decreasing near ω_p^*. The product of a rising and a falling function gives a peak which will be near ω_p^*. We conclude that we expect a large drop to have a peak in the absorption near ω_p^*, compared to a peak at $\omega_p^*/\sqrt{3}$ for a small drop. The large drop should have its peak shifted to higher frequency compared to a small drop of the same density.

The interband contribution in the γ-drop case is much more difficult to compute than for α drops. This is due to the effect of strain on the band structure of Ge. The light hole band and the heavy hole band are split by the strain and at typical values of strain present in the γ drop are both highly aspheric. At the present time, calculations of the interband terms have not been completed. In order to include some interband effects in the lineshape calculation, we have assumed that the interband contribution will be qualitatively similar to the contribution in the absence of strain. We may justify this rather naive assumption by claiming that although the details of the interband transitions will be quite different in the presence of strain, the k-space integrals will "wash out" most of the differences.

In order to calculate an absorption, we need to know the plasma density of a γ-drop. Theoretical calculations[23,24,26,42] suggest that the density in a γ-drop is between 1×10^{16} cm^{-3} and 7×10^{16} cm^{-3},

depending on the amount of strain. The near infrared luminescence linewidth is an experimental estimate of the Fermi energy[28] and hence the density. Using hole effective masses appropriate to unstrained Ge, a density of 6.5×10^{16} cm^{-3} is obtained. However careful consideration of strain effects on the hole masses[49] suggests that this density may be high. At this time the density of a γ-drop ($\langle 110 \rangle$ stress) is not known, but is at least a factor of three lower than the density of an α-drop. Although ω_p^* for a γ-drop is at least $\sqrt{3}$ lower that for an α-drop, the peak γ-drop absorption occurs at ω_p^* rather than $\omega_p^*/\sqrt{3}$. Hence it is not too surprising that the observed absorption maxima occur at the same energy.

We have calculated the absorption of a 100μ radius EHD using Eq. (99) including both plasma and interband contributions to the dielectric constant. The interband terms were those appropriate for unstrained Ge. The dashed line in Fig. 27 is the result of the calculation for $n = 3.6\times10^{16}$ cm^{-3} and $\tau = 8\times10^{-13}$ sec. These parameters were chosen to provide a good fit to the data. Because of the crudeness of the model we are unable to conclude that these values of n and τ are realistic estimates of the actual drop parameters. Theoretical investigations of the γ-drop lineshape are continuing. In order to determine the experimental γ-drop density from our data we must replace Eq. (99) with a full Mie theory calculation and include a more accurate interband dielectric function into the calculation.

0 0 0 4 3 0 8 0 7 0

In spite of the calculation above, the similarity between the γ-drop and the α-drop lineshapes forces a reappraisal of our understanding of the properties of γ-drops. In particular since a large object is responsible for the absorption, we must consider the alternate interpretation that the object is a dense cloud of α-drops rather than a single γ-drop. The first step in the chain of arguments concerning the existence of a γ-drop is the fact that electron-hole drops can be bound in the presence of strain. Suppose we strain the crystal along $\langle 111 \rangle$. Then it is well known[43] that one of the conduction band minima shifts to lower energy while the other three minima shift to higher energy. Theoretical calculations[23-26] and experimental observations of near infrared luminescence[4,44] and the cohesive energy[45] of α-drops under uniaxial stress show that the lifting of the conduction band degeneracy reduces the binding energy of the drop. This reduction in binding energy decreases the density. The two best pieces of experimental evidence for a reduced drop density are the decrease in luminescence signal and the decrease in the luminescence linewidth with increasing stress. The luminescence intensity is proportional to the density. The linewidth is a measure of the Fermi energy which is proportional to $n^{2/3}$. The linewidth is a more persuasive argument for a reduced density because the luminescence intensity will also depend on the relative contributions of radiative and non-radiative decay, but both pieces of evidence support the density reduction.

In a non-uniform stress, α-drops move along the stress gradient toward the maximum stress point and collect there, coalescing into a large γ-drop. The lifetime of the drop is increased because the radiative and non-radiative lifetimes both increase with decreasing density. The radiative lifetime is proportional to n^{-1} while the non-radiative lifetime is proportional to n^{-2} or n^{-3}, depending on the dominant mechanism.

Also the boil-off of excitons from the surface of the drop is inhibited. An exciton which evaporates from the surface of the drop is subjected to the strain gradient which forces it to return to the drop in a time short compared to its lifetime. This is true even though the evaporation rate of excitons from the surface of α-drops is increased under uniaxial stress. The luminescence decay of γ-drops at 1.8K shows no signs of a non-exponential decay due to boiloff. This argues in favor of a single drop rather than a cloud. Suppose we consider an α-drop cloud at the stress maximum. Because of the large surface area, the boiloff will be substantial even at 1.8K. The question of whether the non-exponential decay due to these excitons will be observed clearly depends on the probability that the exciton will collide with an α-drop in a time short compared to its lifetime. This in turn depends on the filling factor of the cloud. Close-packed spheres have a filling factor of .74, but it is unlikely that a cloud could have such a large filling factor because the system could lower its total energy by coalescing into a single drop, thus lowering the surface energy. Considering the effect of the surface energy it seems likely that a cloud would have a filling factor sufficiently low that

1 7 0 8 0 3 4 0 0 0

evaporation effects could be observed. In unstrained Ge, α clouds have typical filling factors of .01 or less.

There is a much stronger argument, however, that the object we call a γ-drop is a single drop rather than a cloud. The microwave dimensional resonances (MDR)[27,28,31] observed for γ-drops require macroscopic circulating currents in the drop. The drop sizes inferred by MDR agree well with sizes measured photographically and by measurements of the luminescence profile of the drop. Therefore currents must circulate all the way out to the surface of the drop. If we assume that the γ-drop is really a cloud, then this current must be due to a hopping conductivity. Estimates indicate that even near the surface of the drop filling factors in the range from .3 to .4 are required to support this current. With such a large filling factor, the surface energy argument would favor coalescence into a single drop. Moreover if such a hopping mechanism were responsible for the MDR signal, estimates indicate that the relaxation time would be of the order of 10^{-15} sec. The actual MDR lines are much too narrow for such a relaxation time. They are well characterized by a relaxation time of 7×10^{-11} sec.

On the basis of the above arguments we have substantial confidence that the γ-drop is a single large electron-hole drop rather than a cloud. We finally must ask, however, whether our absorption be explained by a cloud of drops. Since the α-drops in the cloud would be small, the absorption peak would occur near $\omega_p^*/\sqrt{3}$. The factor of three or more decrease in density, however, would shift the peak lower

in frequency. Thus for an α-drop cloud in stressed Ge we would expect the peak of the absorption to occur below 45 cm^{-1}. This is incompatable with experiment. We must conclude that our observed plasma absorption is due to a single large drop. Unless the experimental and theoretical evidence for the reduction of drop plasma density with stress can be shown to be in error, our results are incompatable with the observation of absorption from a cloud of drops.

F. Magneto-Plasma Effects in Gamma Drops

In our sample geometry (Faraday geometry) it was possible to apply a dc magnetic field of up to 50 kG in a direction normal to the [100] face of the crystal. As is shown in Fig. 29(a), the position of the absorption maximum increased proportional to the square of the applied field. No additional structure was observed in the absorption of the drop, although the full width at half peak absorption coefficient changed in a complex way. The peak absorption coefficient increased with field until approximately 25 kG and then saturated (Fig. 29(b)). The variation of total integrated absorption with field was measured using a broad band technique and the results are shown in Fig. 29(c). These results are similar to the results of Fig. 29(b) but the increase in absorption with field is not as large. This is presumably due to a linewidth effect. In Fig. 29(c) the closed circles were measured while increasing the field while the open circles were measured while decreasing the field. The systematic difference between the measurements in these two directions is thought to be a systematic error in the measurement

0 0 0 0 4 3 0 8 0 7 2

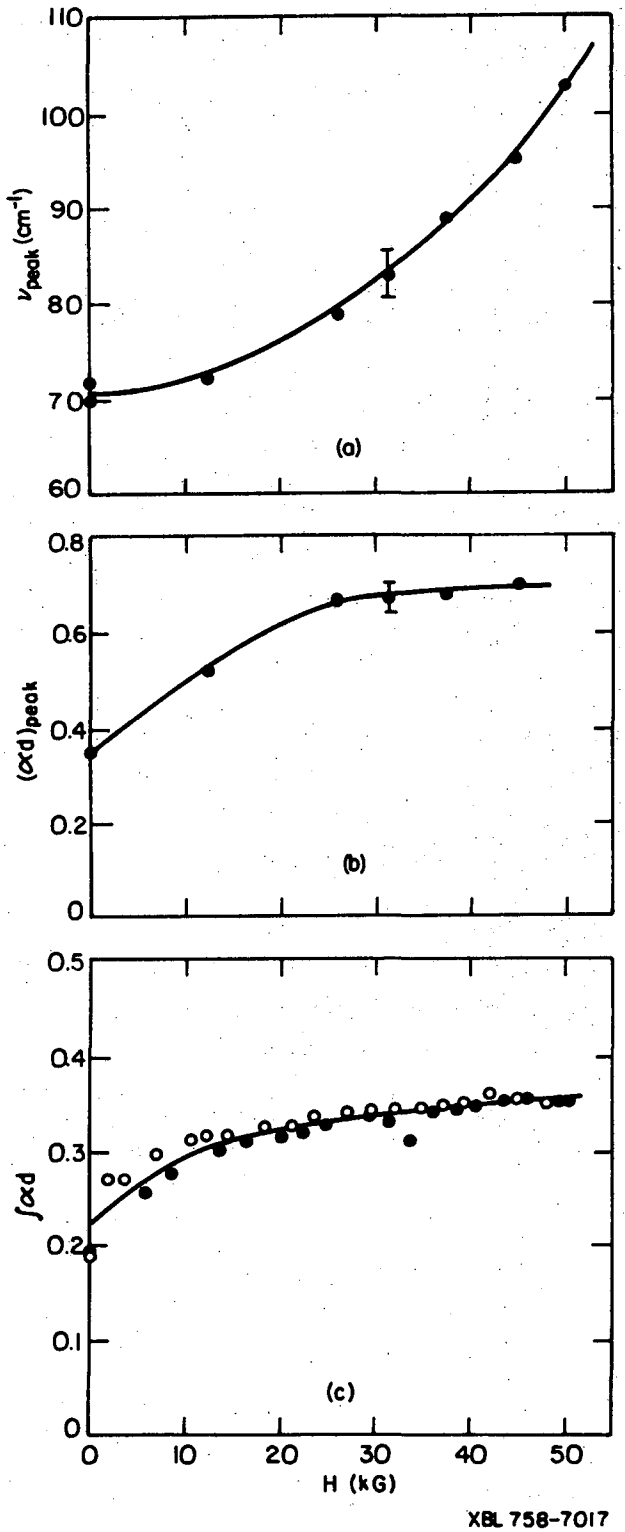


Fig. 29. γ-Drop Magnetoplasma Data.

rather than a real hysteresis effect, but more work is needed to determine this. At the present time we are only prepared to say that no sharp structure was observed in the total absorption. Small variations of the order of one per cent may be present, but the experiment needs to be repeated before more definite conclusions may be drawn.

These results contrast rather sharply with the results of Murzin, et al.,[19,20] Sanada, et al.,[21] and Fujii and Otsuka.[22] In these measurements of α-drops, the resonance at $\omega_p^*/\sqrt{3}$ was observed to split into two components, linearly at low field and more strongly at higher field. The intensity of the absorption (measured either by the peak absorption or by the integrated absorption) decreased markedly with increasing field in the low field limit. A 30% decrease in absorption between zero field and 2 kG was typical. At higher fields the intensity displayed oscillations. The fields at which these oscillations occurred correlated well with the fields at which Landau levels for the electrons cross the electron Fermi energy.

Our interpretation of these data is not complete, so that detailed arguments cannot yet be given. The starting point for consideration of the magnetoplasma absorption is the dielectric constant, which may be written

$$\varepsilon = 1 - \sum_{j=1}^{3} \frac{(\omega_{pj}^*)^2}{\omega(\omega+i/\tau\pm\omega_{cj})} , \qquad (102)$$

0 0 0 0 4 3 0 8 0 7 3

where we have separated the bulk plasma frequency used previously into a sum over electron, light hole, and heavy hole components. The magnetic field dependence enters through the cyclotron frequency for each component,

$$\omega_{cj} = \frac{e_j H}{m_j^* c} \quad , \tag{103}$$

where m_j^* is the cyclotron mass appropriate for component j and $e_j = \pm e$ depending on whether the component j is due to holes or electrons. The cyclotron and plasma masses for the holes are equal, but they are different for electrons.[21] This dielectric function was used by Murzin, et al. to study the dependence of peak position on field for α-drops. His calculations agreed well with experiment and predicted a splitting of the peak. At the present time we do not know why we have not observed a similar splitting. The most likely explantion is that because of size effects the splitting is too small to be resolved. Calculation are underway to verify this idea. Recent calculations by Büttner[50] and Chui[51] show that in strong magnetic fields, the binding energy and hence the density of the drop are field dependent. The magnitude of this effect has not been calculated for the range of fields accessible to our measurements, but it may be an important contribution. No attempt has yet been made to calculate the field dependence of the density of a drop in the low field limit.

At present we have not developed a theory for the field dependence of the total absorption. The most surprising aspect of our results is that no Landau level crossing effects have been observed. This may be due to the large strain inhomogeneities in the drop.

G. Future Experiments

As should be clear from the preceeding discussion, our measurements of electron-hole drops are still in progress and much work remains to be done. In addition to the obvious experiments to improve the data we have already obtained, a number of other experiments suggest themselves.

With regards to α-drops, four major questions seem to me to remain. First we see in Fig. 24 that the high frequency tails of the resonance as shown by four different measurements do not agree. Two possibilities come to mind. The nature of α-drop nucleation centers may effect the local band structure at the nucleation site, thus effecting the interband contribution to the absorption. Therefore the tail of the absorption may vary from sample to sample. On the other hand the absorption could be due to differences in pumping conditions. A focussed spot illumination, for example, would favor formation of an α-cloud. The frequency of the pump light determines the penetration depth. Perhaps there is a difference between the absorption from drops near the surface and drops deep inside the crystal.

Second, the emission from α-drops has never been studied carefully and deserves attention. Vavilov[15] shows that the emission and absorption are roughly similar, but as pointed out by Pokrovskii,[2] Vavilov's

0 0 0 4 3 0 8 0 7 4

interpretation of the origin of the emission is certainly wrong. Both careful experimental and theoretical effort is required.

Third, the photoconductivity experiments described earlier were dropped at an early stage and we have not yet been able to return to them. It is clear that these measurements should be persued.

Finally the absorption of α-drops has never been measured in the presence of uniaxial stress. This experiment should provide important confirmation of the variation of density with stress. The plasma frequency should shift proportional to $n^{1/2}$. The shift in the interband term is harder to estimate, but the limits of integration of the interband integral I (Eq. (92)) shift proportional to $n^{1/3}$.

For the γ-drop experiments, the most important experiment is to try different samples. Since only one sample was used in the γ-drop work, we have no idea at present of whether any of our results on sample dependent. Also we have observed large offset signals from γ-drops when we have chopped the pump light rather than the interferometer output. This signal is probably due to γ-drop emission and should be studied. Finally the magnetoplasma work should be expanded to include other sample orientations.

CHAPTER VI REFERENCES

1. V. M. Asnin, A. A. Rogachev, and S. M. Ryvkin, Sov. Phys. Semicond. 1, 1447 (1968); A. A. Rogachev, Proc. 9th Int. Conf. Phys. Semiconductor, Moscow (1968), p. 407; and V. M. Asnin and A. A. Rogachev, JETP Letters 14, 338 (1971).
2. Ya. Pokrovskii, Phys. Stat. Sol. (a) 11, 385 (1972).
3. Ya. Pokrovskii and K. I. Svistunova, JETP Letters 9, 261 (1969).
4. C. Benoit a la Guillaume, M. Voos, and F. Salvan, Phys. Rev. B5 3079 (1972).
5. C. Benoit a la Guillaume and M. Voos, Phys. Rev. B 7, 1723 (1973).
6. T. K. Lo, B. J. Feldman, and C. D. Jeffries, Phys. Rev. Letters 31, 224 (1973).
7. G. A. Thomas, T. G. Phillips, T. M. Rice, and J. C. Hensel, Phys. Rev. Letters 31, 386 (1973).
8. T. K. Lo, Solid State Commun. 15, 1231 (1974).
9. G. A. Thomas, T. M. Rice, and J. C. Hensel, Phys. Rev. Letters 33, 219 (1974).
10. T. K. Lo, B. J. Feldman, R. M. Westervelt, J. L. Staehli, and C. D. Jeffries, to be published.
11. J. C. Hensel, T. G. Phillips, and T. M. Rice, Phys. Rev. Letters 30, 227 (1973).
12. R. M. Westervelt, T. K. Lo, J. L. Staehli, and C. D. Jeffries, Phys. Rev. Letters 32, 1051 (1974).
13. Ya. Pokrovskii and K. I. Svistunova, Sov. Phys. Semicond. 4, 409 (1970).

0 0 0 0 4 3 0 8 0 7 5

14. J. M. Worlock, T. C. Damen, K. L. Shaklee, and J. P. Gordon, Phys. Rev. Letters 33, 771 (1974).

15. V. S. Vavilov, V. A. Zayats, and V. N. Murzin, JETP Letters 10 192 (1969).

16. V. S. Vavilov, V. A. Zayats, and V. N. Murzin, Proc. 10th Int. Conf. Phys. Semicond., Cambridge, Mass (1970), p. 509.

17. V. L. Kononenko, T. S. Mandelstam, V. N. Murzin, V. S. Vavilov, and V. A. Zayats, Proc. 12th Int. Conf. Phys. Semiconductors, Stuttgart (1974), p.152.

18. T. Timusk and A. Silin, to be published.

19. V. N. Murzin, V. A. Zayats, and V. L. Kononenko, Proc. 11th Int. Conf. Phys. Semiconductors, Warsaw (1972), p. 678.

20. V. N. Murzin, V. A. Zayats, V. L. Kononenko, Sov. Phys. Solid State 15, 2421 (1974).

21. T. Sanada, T. Ohyama, and E. Otsuka, Solid State Commun. 12, 1201 (1973).

22. K. Fujii and E. Otsuka, to be published.

23. M. Combescot and P. Nozieres, J. Phys. C: Solid State Phys. 5, 2369 (1972).

24. W. F. Brinkman and T. M. Rice, Phys. Rev. B 7, 1508 (1973).

25. P. Vashista, P. Bhattacharyya, and K. S. Singwi, Phys. Rev. Letters 30, 1248 (1973).

26. P. Vashista, S. G. Das, and K. S. Singwi, Phys. Rev. Letters 33, 911 (1974).

27. R. S. Markiewicz, J. P. Wolfe, and C. D. Jeffries, Phys. Rev. Letters, 32, 1357 (1974).

28. J. P. Wolfe, R. S. Markiewicz, C. Kittel, and C. D. Jeffries, Phys. Rev. Letters 34, 275 (1975).

29. J. P. Wolfe, W. L. Hansen, E. E. Haller, R. S. Markiewicz, C. Kittel, and C. D. Jeffries, Phys. Rev. Letters 34, 1292 (1975).

30. J. P. Wolfe, S. M. Kelso, R. S. Markiewicz, and J. E. Furneaux, to be published.

31. R. S. Markiewicz, Phys. Rev. B 10, 1766 (1974).

32. H. Hinterberger and R. Winston, Rev. Sci. Instr. 37, 1094 (1966); R. Winston, J. Opt. Soc. Am. 60, 245 (1970).

33. Our Ge samples typically contained less than 3×10^{10} ionizable impurities per cm^3. They were provided by W. L. Hansen and E. E. Haller.

34. G. Mie, Ann. d. Physik (4) 25, 377 (1908).

35. M. Born and E. Wolf, Principles of Optics, (Pergamon Press, Oxford, 1965), Section 13.5.

36. J. D. Jackson, Classical Electrodynamics, (John Wiley and Sons, New York, 1967), Section 16.9.

37. H. C. van de Hulst, Optics of Spherical Particles, (N. V. Drukkerij, J. F. Duwaer & Zonen, Amsterdam 1946), Ph.D. Thesis.

38. H. C. van de Hulst, Light Scattering by Small Particles, (John Wiley & Sons, New York, 1957).

39. T. M. Rice, Bull. Am. Phys. Soc. 20, 470 (1975).

40. J. Rose and H. B. Shore, Bull Am. Phys. Soc. 20, 470 (1975).

41. M. Combescot and P. Nozieres, Solid State Commun. 10, 301 (1972).

42. P. Vashista, private communication.

0 0 0 0 4 3 0 8 0 7 6

43. I. Balslev, Phys. Rev. 143, 636 (1966).

44. A. S. Alekseev, V. S. Bagaev, and T. I. Galkina, JETP 36, 536 (1973).

45. T. Ohyama, T. Sanada, and E. Otsuka, Phys. Rev. Letters 33, 647 (1974).

46. P. Debye, Ann. d. Physik (4) 30, 118 (1909).

47. T. L. Reinecke and S. C. Yang, Phys. Rev. Letters 35, 311 (1975).

48. J. P. Wolfe, R. S. Markiewicz, and C. D. Jeffries, Proc. 3rd Int. Conf. Light Scattering in Solids, Campinas, Brasil (1975), to be published.

49. R. S. Markiewicz, private communication.

50. H. Büttner, Proc. 12th Int. Conf. Phys. Semiconductors, Stuttgart (1974).

51. S. T. Chui, Phys. Rev. B 9, 3438 (1974).

ACKNOWLEDGEMENTS

It is with pleasure and appreciation that I thank the many people whose contributions of time and energy made this work possible. I would first like to thank my advisor, P. L. Richards for his interest, insight, and support. When I began this work I was a person who had studied some physics. If today I occasionally think like a physicist, he is responsible. I would also like to thank my fellow graduate students John Mather and David Woody who were always willing to rally 'round when a problem arose. The personnel of the Physics Department mechanical and electronic shops deserve recognition for their assistance in the fabrication of the instrumentation developed in this work.

I would like to gratefully acknowledge very useful theoretical discussions about the $Al_2O_3:Mn^{3+}$ and γ-irradiated ruby experiments with C. A. Bates and L. Falicov. I also had extensive discussions with R. A. Forman, who shared with me the results of unpublished measurements on both systems. I am indebted to R. A. Forman and B. Suffer for providing samples and to G. Shalimoff for the spectrochemical analysis. Norman Nishioka provided assistance with the computer programming.

The electron-hole drop work was performed in partnership with Laurence Eaves. I would also like to thank C. D. Jeffries, Robert Markiewicz, and James Wolfe for helpful discussions. Samples for this work were provided by Eugene Haller.

I would like to thank Alice Ramirez for typing the manuscript and Gloria Pelatowski for drawing the figures.

0 0 0 4 3 0 8 0 7 7

Most of all I would like to thank my wife. Her willingness to share the disappointment of my failures along with the excitement of my successes, her constant enthusiasm, support, encouragement, and understanding truly made this work possible.

This work was supported in part by the U. S. Energy Research and Development Administration.

APPENDIX A.

Computer I/O Register Assignments

In this appendix, the peripherals which have been interfaced to the PDP 11 are described from the viewpoint of the programmer who wishes to use them in a program. For each peripheral, a brief description is included of what it does. The peripheral registers assigned to it are listed, as well as the vector addresses and priorities of any interrupts generated by the device. Each peripheral register is then described in detail using a register map and a functional description of each bit. There are four ways that a bit may be implemented in a peripheral register. It may be a read/write bit (RW) in which case it may be read or written by the computer. Read/write bits are normally used to transfer information to the peripheral. It may be read-only (RO). This type of bit is used to transfer information from the peripheral to the computer. Writing it has no effect. A bit may be write-only (WO). In this case the bit will always read as a zero. Write-only bits are generally used as commands to the device to initiate action. Finally the bit may be unimplimented. It will always read as a zero. Writing it will not convey information to the device.

As well as the custom-designed peripherals, information is included on the standard DEC peripherals interfaced to the computer. Information concerning these devices was obtained from handbooks published by DEC. The reader's attention is in particular directed toward the PDP 11 Peripherals Handbook which contains descriptions of available standard peripherals and a thorough discussion of the theory

0 0 0 0 4 3 0 8 0 7 8

and operation of the Unibus and Unibus interfacing.

1. Teletype, KL11

The teletype is a standard ASR33 Model made by Teletype Corporation. It transmits and receiver data serially at a rate of 10 characters/second and consists of four parts, a keyboard, printer, paper tape reader and papertape punch. The reader and punch operate in parallel with the keyboard and printer respectively and may be turned on and off by means of special control characters.

Registers:

Name	Description	Address
TKS	Teletype Keyboard Status	177560
TKB	Teletype Keyboard Buffer	177562
TPS	Teletype Printer Status	177564
TPB	Teletype Printer Buffer	177566

Interrupts:

Name	Vector Address	Priority
Keyboard Done	60	4
Printer Ready	64	4

TKS 177560

				Busy				Done	Intr Enb						Rdr Enb
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0

bit 0 -- (WO) Reader Enable. Set to allow the papertape reader (not the keyboard) to read one character.

bits 1-5 Not used.

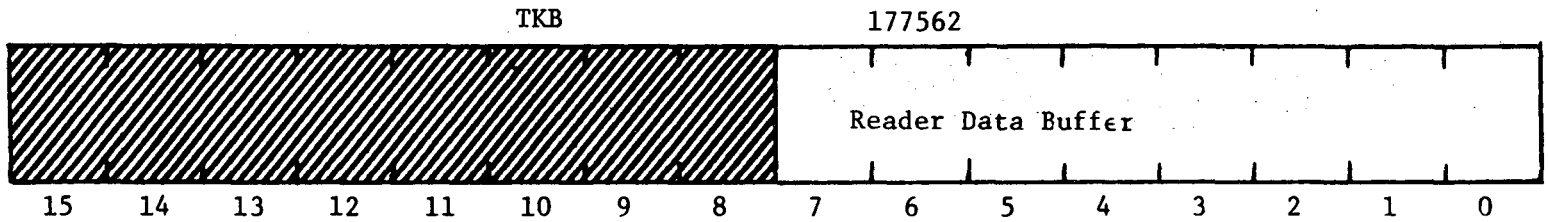
bit 6 -- (RW) Interrupt Enable. Set to allow Done = 1 to cause an interrupt.

bit 7 -- (RO) Done. Set when a character is available in the Reader Data Buffer. It is cleared by any program reference to the Reader Data Buffer, or when Reader Enable is set.

bits 8-10 Not used.

bit 11 (RO) Busy. Set when receiving information. It is set by the detection of the beginning of a character and is cleared by the leading edge of Done.

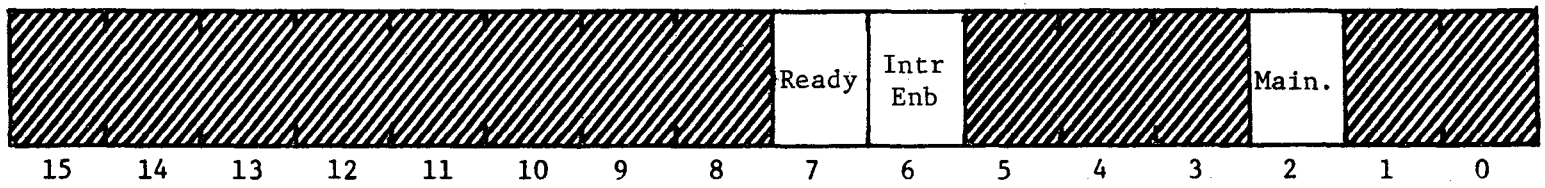
0 0 0 0 4 3 0 8 0 7 9



bits 0-7 (RO) Reader Data Buffer. These bits hold the coded data for the character read. Any program reference to TKB will clear Done, bit 7 of TKS.

bits 8-15 Not used.

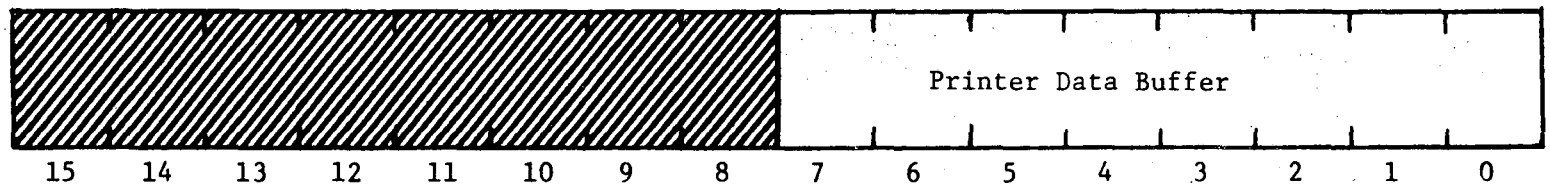
TPS 177564



bits 0-1	Not used.
bit 2	(RW) Maintenance. Set to disable the serial line input from the teletype and enable serial output of the printer to feed into the reader. It is used for maintenance purposes.
bit 3-5	Not used.
bit 6	(RW) Interrupt Enable. Set to allow Ready = 1 to cause an interrupt.
bit 7	(RO) Ready. Set when the printer is available to accept data. It is cleared when data is loaded into the Printer Data Buffer and is set when another character can be loaded.

0 0 0 0 4 3 0 8 0 8 0

TPB 177566



bits 0-7 (WO) Printer Data Buffer. These bits hold the coded data for the character to be printed. Loading the register initiates the character printing cycle. Any instruction that could modify TPB clears Ready, bit 7 of TPS.

bits 8-15 Not used.

2. Papertape Reader/Punch, PC11

The reader reads unoiled, fan-folded, 8-hole perforated papertape at 300 characters per second using photoelectric detection of the holes in the tape. Synchronization of the read operation is maintained by optically detecting the sprocket hole. The punch will perforate tape at the rate of 50 characters per second.

Registers:

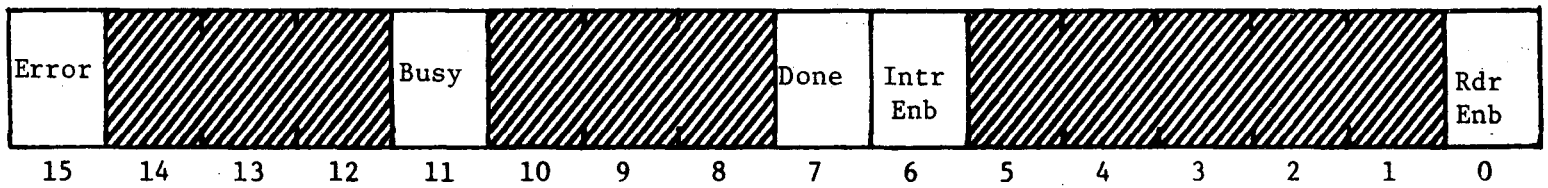
Name	Description	Address
PRS	Papertape Reader Status	177550
PRB	Papertape Reader Buffer	177552
PPS	Papertape Punch Status	177554
PPB	Papertape Punch Buffer	177556

Interrupts:

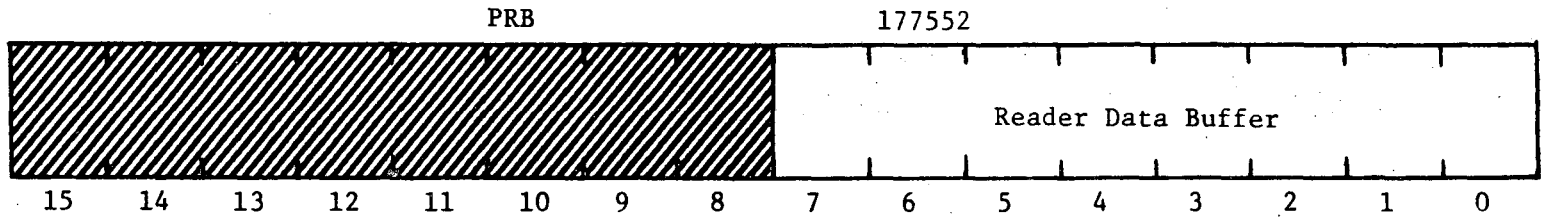
Name	Vector Address	Priority
Reader Done or Error	70	4
Punch Ready or Error	74	4

0 0 0 0 4 3 0 8 0 8 1

PRS 177550



bit 0	(WO) Reader Enable. Set to allow the reader to fetch one character. The setting of this bit clears Done, sets Busy, and clears PRB. If Error = 1, setting this bit will cause an interrupt if Intr Enb = 1, and reading will be inhibited.
bits 1-5	Not used.
bit 6	(RW) Interrupt Enable. Set to allow Done or Error = 1 to cause an interrupt.
bit 7	(RO) Done. Set when a character is available in PRB. It is cleared by any program reference to PRB or by setting Reader Enable.
bits 8-10	Not used.
bit 11	(RO) Busy. Set when a character is being read. Set by Reader Enable and cleared by Done.
bits 12-14	Not used.
bit 15	(RO) Error. Set when an error occurs; no tape in reader, reader off-line, or no reader power.

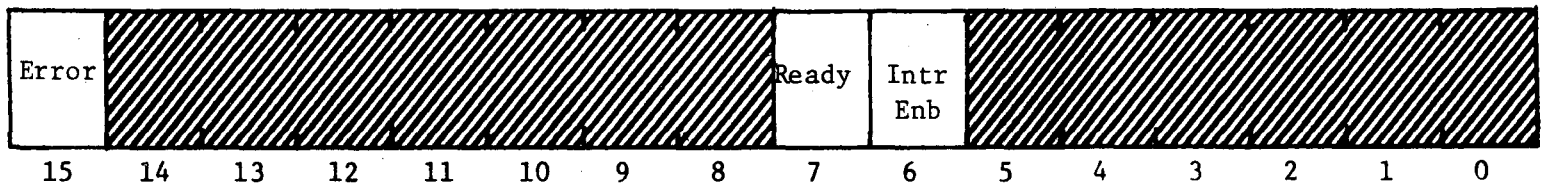


bits 0-7 (RO) Reader Data Buffer. These bits hold the coded data for the character read. These bits are cleared when Reader Enable, bit 0 of PRS is set. Any program reference to PRB will clear Done, bit 7 of PRS.

bits 8-15 Not used.

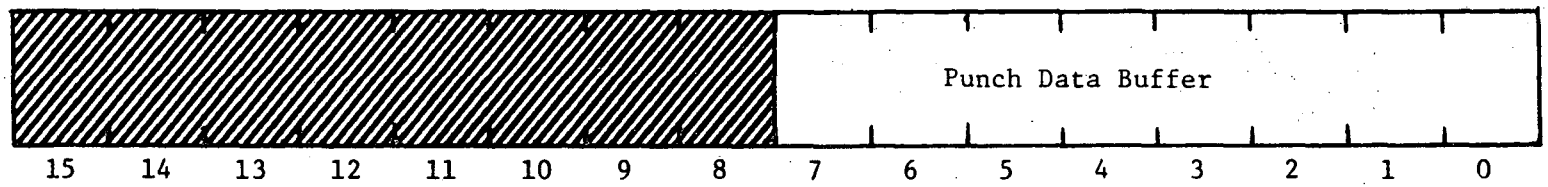
0 0 1 0 4 3 0 8 0 8 2

PPS 177554



bits 0-5	Not used.
bit 6	(RW) Interrupt Enable. Set to allow Error or Ready = 1 to cause an interrupt.
bit 7	(RO) Ready. Set when ready to punch a character. It is cleared when PPB is loaded, and is set when punching is done.
bits 8-14	Not used.
bit 15	(RO) Error. Set when an error occurs; no tape in punch, or punch has no power.

PPB 177556



bits 0-7 (WO) Punch Data Buffer. These bits hold the coded data for the character to be punched. Any instruction that could modify PPB clears Ready, bit 7 of PPS, and initiates punching. An immediate interrupt will occur when punching is initiated if Error = 1 and Interrupt Enable = 1.

bits 8-15 Not used.

0 0 0 0 4 3 0 8 0 8 3

3. Line Frequency Clock, KW11-L

The line clock provides timing information for the computer using a 60 Hz clock derived from the ac line. In interrupt mode, an interrupt is generated each cycle of the line frequency. In non-interrupt mode the program may check the Monitor bit for timing information.

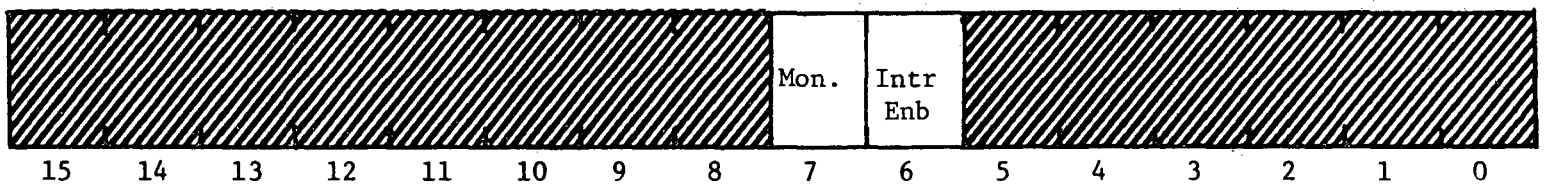
Register:

Name	Description	Address
LKS	Line Clock Status	177546

Interrupt:

Name	Vector Address	Priority
Monitor	100	6

LKS 177546



bits 0-5 Not used.

bit 6 (RW) Interrupt Enable. Set to allow Monitor = 1 to cause an interrupt.

bit 7 (RW) Monitor. Set by the line frequency clock signal and cleared by the program.

0 0 0 0 4 3 0 8 0 8 4

4. Programmable Real-Time Clock, KW11-P

This device provides interrupts at programmable intervals. It consists of a 16-bit counter that counts up or down at four selectable rates. These rates are 100 kHz, 10 kHz, 60 Hz and a frequency determined by an external input. The external input is provided by the clock section of the DAS. When counting down, the interrupt is initiated at zero (underflow). The up-count mode is used for event counting and timing. An interrupt is initiated at zero (Overflow). A count set register stores the preset count interval. In repeat interrupt mode, it reloads the counter on underflow. The control and status register controls operation of the unit.

In single-interrupt mode, an interrupt is generated at the end of the preset interval. The clock is stopped and the counter is set to zero. In repeat interrupt mode, the counter is automatically reloaded every time an interrupt occurs, and clocking is restarted.

Register:

Name	Description	Address
PKS	Clock Status and Control	172540
PKB	Count Set Buffer	172542
PKC	Counter Register	172544

Interrupt:

Name	Vector Address	Priority
Interval Done	104	6

PKS 172540

Error								Done	Intr Enb	Fix	Up/Dn	Mode	Rate Select		Run
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0

bit 0 (RW) Run. Set to allow the counter to count. Cleared on underflow or overflow in Single-Interrupt Mode.

bits 1-2 (RW) Rate Select. Selects one of four available clock rates.
00 - 100 kHz; 01 - 10 kHz; 10 - 60 Hz; 11 - External.

bit 3 (RW) Mode. Selects interrupt mode of operation. 1 = repeated interrupts, 0 = single interrupt.

bit 4 (RW) Up/Down. If = 1, counter counts up. If = 0, counter counts down.

bit 5 (WO) Fix. Maintenance bit. Causes single clocking of the counter.

bit 7 (RO) Done. Set on counter overflow or underflow.

bit 8-14 Not used.

bit 15 (RO) Error. In repeat interrupt mode, a second overflow or underflow before the first is serviced sets this bit. It is cleared by addressing PKS.

0 0 0 4 3 0 8 0 8 5

PKB 172542

Count Set Buffer Register															
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0

bits 0-15 (WO) This 16-bit register is used for storage of the interval count. It allows automatic reloading of the counter in repeat-interrupt mode. In single-interrupt mode, an underflow or overflow clears the buffer.

PKC 172544

Counter Register															
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0

bits 0-15 (RO) This 16-bit register is a binary up/down counter. It is the counter controlled by PKS. Overflow or underflow of this counter sets Done.

5. Extended Arithmetic Element (EAE), KE11-A

This device performs signed integer multiplication and division, arithmetic and logical multiple position shifts, and normalization at speeds much faster than is possible using software routines. It multiplies two 16-bit signed integers to give an 32-bit product and divides a 32-bit dividend by a 16-bit divisor to give a 16-bit quotient and a 16-bit remainder. A 32-bit number can be shifted either left or right the number of places specified by a count (logical shift). In an arithmetic shift, the sign of the number is prevented from changing. In the normalization operation, a 32-bit number is shifted left until the two most significant bits are different. A count is kept of the number of places the number is shifted. This operation is useful for programming floating-point operations. For a more complete description of EAE, the reader is referred to the KE11-A Users Manual, published by DEC.

Registers:

Name	Description	Address
DIV	Divisor	177300
AC	Accumulator	177302
MQ	Multiplier/Quotient	177304
MUL	Multiplicant	177306
SC	Step Count	177310
SR	Status Register	177311
NORM	Normalize	177312
LSH	Logical Shift	177314
ASH	Arithmatic Shift	177316

6. Data Acquisition System

The data acquisition system has been described in detail in Chapter III.

Registers:

Name	Description	Address
SSN	Slo-Syn States Register	167770
DSP	Display Register	167772
ADB	ADC Data Register	167774
ADC	ADC Status Register	167776

Interrupts:

Name	Vector Address	Priority
Slo-Syn Done	150	6
External Interrupt	154	6
ADC Done	160	6

SSN 167770

Ext Intr Flag					Dir.		Jog	Slo-Syn Done	Slo-Syn Intr Enb	Ext Intr Enb					Index
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0

bit 0	(WO) Index. When set, issues a command to index the Slo-Syn.
bits 1-4	Not used.
bit 5	External Interrupt Enable. When set, External Interrupt Flag = 1 will cause an interrupt.
bit 6	(RW) Slo-Syn Interrupt Enable. When set, Slo-Syn Done = 1 will cause an interrupt.
bit 7	(RO) Slo-Syn Done. Set by completion of motor movement. Cleared by Index or Jog.
bit 8	(WO) Jog. When set, issues a command to jog the Slo-Syn.
bit 9	Not Used.
bit 10	(RW) Direction. Motor moves forward when clear, in reverse when set.
bits 11-14	Not used.
bit 15	(RO) External Interrupt Flag. This bit is set except for a microsecond whenever an external interrupt request is detected.

0 0 0 0 4 3 0 8 0 8 7

DSP 167772

Display Register															
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0

bits 0-15 (RW) Display Register. The contents of this register (considered a 16-bit two's complement binary number) is converted to a signed decimal number and displayed on the DAS readout.

ADB 167774

Data Register															
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0

bits 0-15 (RO) Data Register. The result of an ADC conversion is placed in this register.

ADC 167776

Over Range Flag	Mon. Flag				Multiplexer Channel			ADC Done	ADC Intr Enb						CTC
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0

bit 0 (WO) Command to convert. Sends convert pulse to DAS when set.

bits 1-5 Not used.

bit 6 (RW) Enable. When set, ADC Done will cause an interrupt.

bit 7 (RO) ADC Done. Set when ADC conversion is complete. Cleared by initiation of a conversion.

bits 8-10 (RW) Multiplexer Channel. When the DAS is set for computer selection, these bits determine the channel. Bit 8 is the LSB.

bits 11-13 Not used.

bit 14 (RW) - Monitor Flag. This bit is a software flag usable for any desired purpose.

bit 15 (RO) - Over Range Flag. When set, it indicates that the previous ADC conversion resulted in an overrange condition.

0 0 0 0 4 3 0 8 0 8 8

7. Oscilloscope

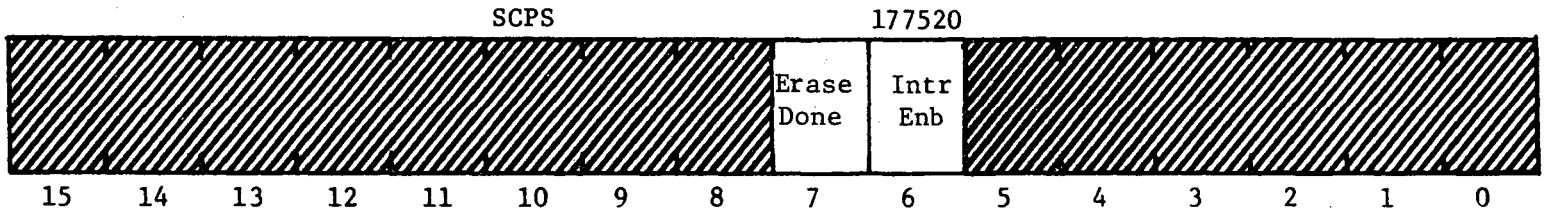
A Tektronix Model 601 Storage Scope is interfaced to the system. A pair of 10-bit unipolar binary DACs provide 0-1V signals to the X and Y inputs of the scope. The scope controller also turns the electron beam on and off, switches the scope between storage and non-storage modes of operation, and erases the screen.

Registers:

Name	Description	Address
SCPS	Scope Status	177520
SCPY	Scope Y-Axis	177522
SCPX	Scope X-Axis	177532

Interrupt:

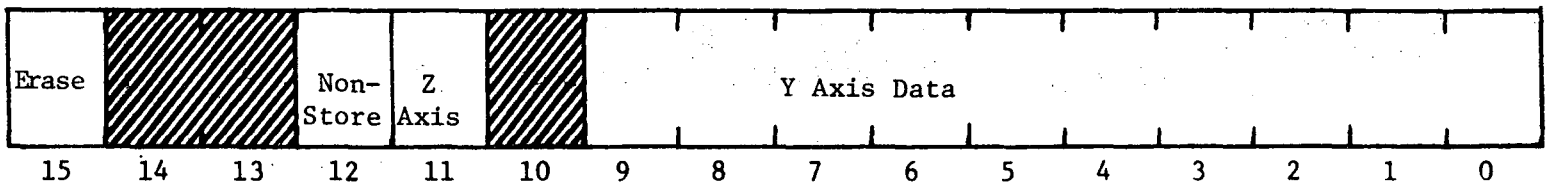
Name	Vector Address	Priority
Erase Done	130	5



bits 0-5	Not used.
bit 6	(RW) Intr Enb. When set, Erase Done = 1 will cause an interrupt.
bit 7	(RO) Erase Done. Cleared by Erase. The bit is asserted when the scope erase cycle is completed.
bits 8-15	Not used.

0 0 0 0 4 3 0 8 0 8 9

SCPY 177522



bits 0-9 (RW) Y Axis Data. The contents of these bits are used to load the unipolar 10-bit binary DAC which is connected to the Y axis of the scope. Loading this register causes both Y and X axis information to be strobed into the respective DACs.

bit 10 Not used.

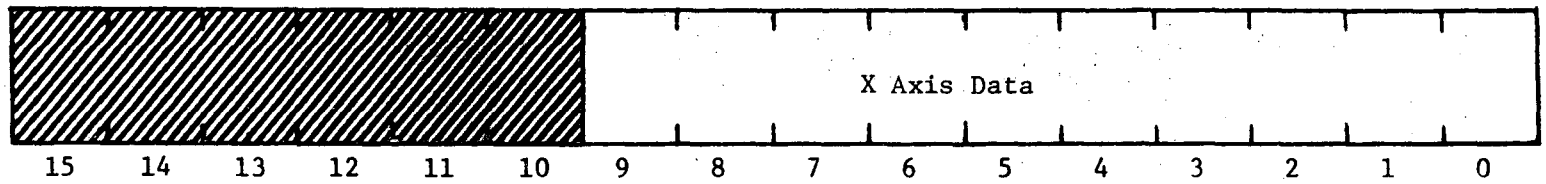
bit 11 (RW) Z Axis. When set, the beam of the scope is turned on.

bit 12 (RW) Non-store. When clear, the scope is placed in storage mode. When set, the scope is placed in non-store mode.

bits 13-14 Not used.

bit 15 (RW) Erase. When set, an erase cycle is initiated.

SCPX 177532



bits 0-9 (RW) X Axis Data. The contents of these bits are used to load the unipolar 10-bit binary DAC which is connected to the X axis of the scope. The contents of these bits are strobed into the DAC whenever the SCPY register is written.

bits 10-15 Not used.

0 0 0 0 4 3 0 8 0 9 0

8. Card Punch

An IBM 526 cardpunch is controlled by a series of high-voltage transistor switches. Relay logic is used for detection of the completion of a punch operation. Characters are converted from 7-bit ASCII code (generated by the computer) to 12-bit Hollerith code (used by the cardpunch) using a programmed Intel 1702A read-only memory. Two versions of the Hollerith code are selectable. The code used by the LBL Computer Center (026 code) is an out-of-date code. The new standard Hollerith code (029 code) is used in IBM 029 keypunches.

Registers:

Name	Description	Address
CPS	Card Punch Status	167400
CPB	Card Punch Buffer	167402

Interrupt:

Name	Vector Address	Priority
Card Punch Ready	164	5

CPS 167400

Error								Ready	Intr Enb				Code Sel.		
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0

bits 0-1 Not used.

bit 2 (RW) Selects between old (026) Hollerith and new (029) Hollerith codes for the character being punched. When clear, 026 code is selected.

bits 3-5 Not used.

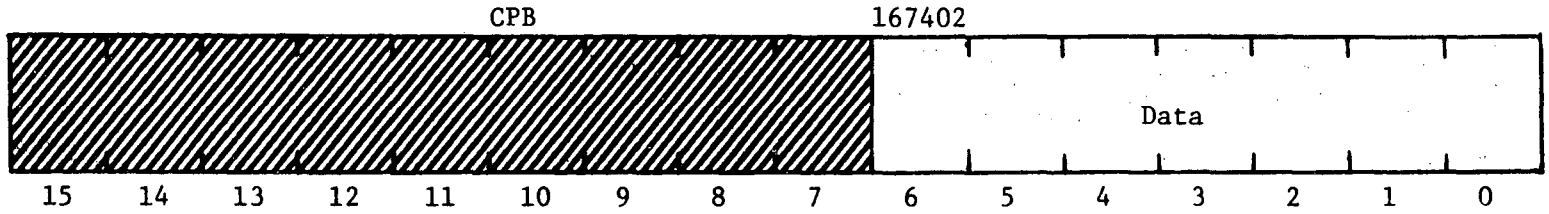
bit 6 (RW) Interrupt Enable. When set, Ready or Error = 1 causes an interrupt.

bit 7 (RO) Ready. Writing CPB clears Ready. The bit is set when the punch has finished punching a character and is ready for the next character. If registration of a new card is required before the punch is ready for the next character, Ready is delayed until this has occurred.

bits 8-14 Not used.

bit 15 (RO) Error. This bit is set if the punch runs out of cards. Error is cleared by the rectification of the error condition.

0 0 0 0 4 3 0 8 0 9 1



bits 0-6 (WO) Data. Writing the 7-bit ASCII code for a character causes the character to be punched

bits 7-15 Not used.

9. X-Y Plotter

A special interface was designed for a Hewlitt - Packard Model 7004B X-Y Plotter. The x- and y- inputs are connected to the outputs of two 12-bit unipolar binary DACs. These DACs are controlled by the outputs of two 12-bit up-down counters. Two modes of operation are possible. The counters may be directly loaded by the computer, causing the pen to move to the new coordinates at its maximum slew rate. In addition, the up and down clocks of the counters are also available. Using them will cause the counters to increment or decrement, moving the pen incrementally. Software has been derived to use this incremental plotting mode for generating vectors which are straight lines between points. The pen is driven at constant velocity in this mode, rather than at a constant time per point. The pen may also be raised and lowered under computer control. A flag is available which indicates when the pen has finished moving, although no interrupt is generated.

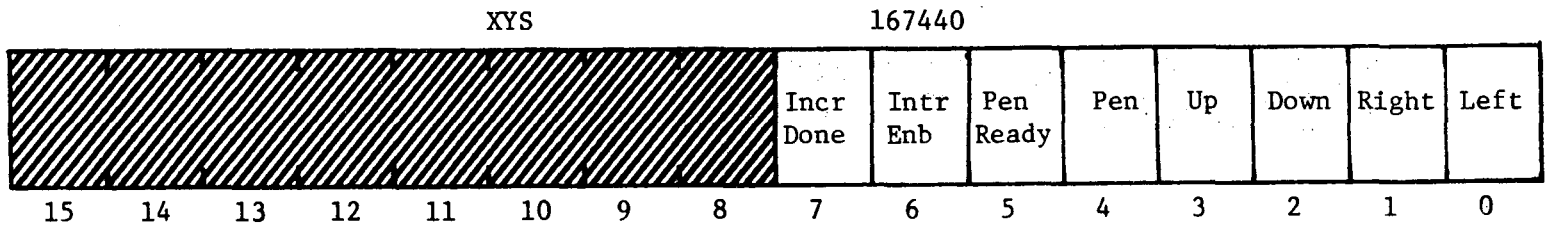
Registers:

Name	Description	Address
XYS	Plotter Status Register	167440
XYY	Y-Axis Position	167442
XYX	X-Axis Position	167444

Interrupt:

Name	Vector Address	Priority
Increment Done	170	5

00043080962



bit 0 (WO) Left. Setting this bit decrements XYX, moving the pen one unit left.

bit 1 (WO) Right. Setting this bit increments XYX, moving the pen one unit right.

bit 2 (WO) Down. Setting this bit decrements XYY, moving the pen one unit down.

bit 3 (WO) Up. Setting this bit increments XYY, moving the pen one unit up.

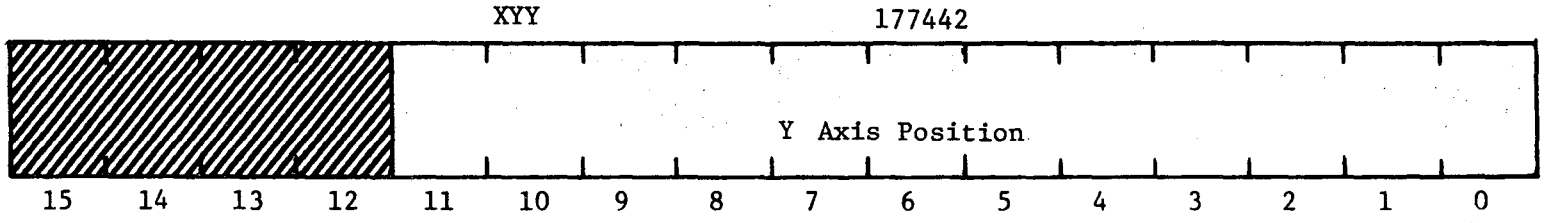
bit 4 (RW) Pen. Setting this bit lowers the pen to the paper. Pen = 0 lifts pen.

bit 5 (RO) Pen Ready. Set when a pen motion is completed. Cleared by changing the value of bit 4, Pen.

bit 6 (RW) Interrupt Enable. If set, Incr Done = 1 will cause an interrupt.

bit 7 (RO) Increment Done. Cleared by writing any of the pen motion bits 0-3. It becomes set when the pen has responded to the increment command.

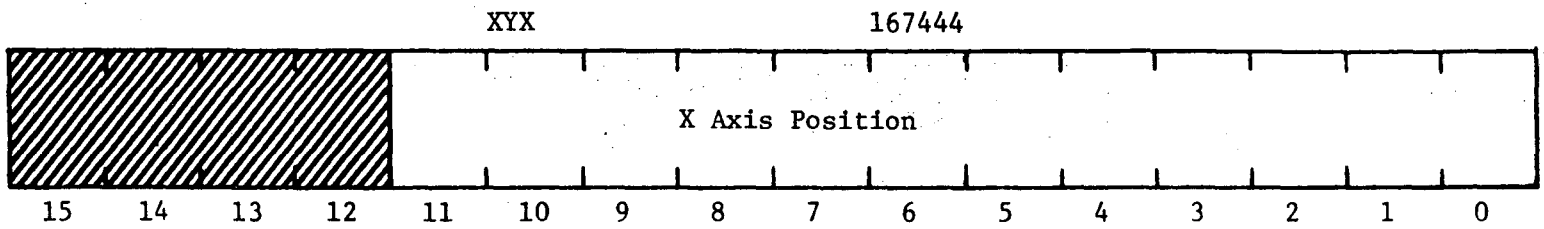
bits 8-15 Not used.



bits 0-11 (RW) Y Axis Position. These bits contain the current y coordinate of the pen position. This register is a 12-bit up/down counter. It may be directly loaded by writing this location. It may be also incremented or decremented by the appropriate bits of XYS. The outputs of the counter are connected to the inputs of a unipolar 12-bit binary DAC connected to the y-axis input of the x-y plotter.

bits 12-15 Not used.

0 0 1 0 4 3 0 8 0 9 3



bits 0-11 (RW) X Axis Position. These bits contain the current x coordinate of the pen position. This register is a 12-bit up/down counter. It may be directly loaded by writing this location. It may also be incremented or decremented by the appropriate bits of XYS. The outputs of the counter are connected to the inputs of the unipolar 12-bit binary DAC connected to the x-axis input of the x-y plotter.

bits 12-15 Not used.

APPENDIX B.

Proposed Scope Controller

A very large number of choices exist when considering the design of a visual display system and in order to understand a particular design, the philosophy which governs the choices made is important. Basically we wish to display graphical information in the form of plots of data points (interferograms, spectra, ratios, etc.) as a function of a horizontal coordinate (frequency or wavenumber). Although a desirable luxury, alphanumeric characters need not be incorporated into the display. Also we do not plan to draw pictures on the screen, so vector generation is not necessary. The display should be capable of rapid change so that the operator could change display scaling easily and quickly. The display software should be simple and require little computer memory. Display processing should requires only a small fraction of available computer time.

The first decision to be made is whether a refreshed or stored display should be used. We opt for a refreshed display because it is easier to change to reflect real-time changes in the data. Next we must decide on whether the refreshed display will be based on a TV monitor, a conventioanl x-y display monitor, or a storage monitor operated in non-storage mode. A TV monitor, which uses a composite video signal, is much better suited to alphanumeric displays than to graphics because of the excessively high memory requirements in a high resolution graphics system. To store a video image based on a 1024 x 1024 raster requires 1,048,576 bits of

0 0 0 4 3 0 8 0 9 4

storage and hence is not feasible for graphics applications. A conventional x-y display monitor is better suited to graphics applications. Since we are already using a storage monitor, we will design our system around this scope. It has the advantage that multiple data sets may be easily displayed in storage mode for easy visual comparison.

In any refreshed display, there is a trade-off between the amount of information that can be placed on the screen and the time required to write that information. If the display is refreshed too slowly, it will appear to flicker. In order to prevent flicker, the data must be refreshed at least every 30 ms. Suppose that we require that a 1024 point display show no objectionable flicker. Then we are allowed only 30 μs to write each point on the screen. This stringent time requirement will affect both implementation of data scaling and the way in which the refresh is done.

In order for the display to be useful, it should be possible for the operator to shift the position of the origin and to change the gain of the display. For example, if some small feature is of interest, the operator may wish to suppress the origin of the display and expand the scale in order to make the feature more prominent. The scaling algorithm will be discussed below. The important point here is that the faster possible software scaling algorithm requires more than 30 μs per point. There are only two alternatives. Either a display buffer must be reserved in memory or the algorithm must be implemented in hardware. If a display buffer were used, the data would be placed in the buffer and scaled wherever a change

occurred in either the data or the sealing. While feasible, this approach would be very wasteful of computer memory. A far superior approach is to scale the data in hardware on a point-by-point basis.

There are basically four ways of providing display refresh: under program control, under interrupt control, with internal display memory, and with direct memory access (DMA). The present display is run under program control. Since this is a background task, no other background tasks can be run concurrently. While it might seem that interrupt control (foreground mode) is the obvious way to implement a display controller, it would be inappropriate in this case. The fastest possible interrupt service routine would require at least 15 μs so that at best background tasks would have only 50% of the available time in which to be executed. While 50% dead time is a significant improvement on the 100% dead time we have at present, better performance can be achieved. The next best alternative is to build an internal display buffer into the controller. The refresh would then take place from this internal buffer and no processor time would be required except to change the contents of the display buffer when the data changed. This option would require the design of semiconductor memory and would probably be too expensive of time and materials to be a good choice. The best option is DMA refresh. In this system, special circuitry is used to refresh the display directly from computer memory without processor intervention. Although extra circuitry is required, most of the added parts have been already documented in other applications, so the amount of new design is fairly small. With this option, background tasks could be executed with only 4% dead time.

0 0 0 4 3 0 8 0 9 5

Control Circuitry

The display controller consists of DMA interface circuits, control logic, x and y axis circuits, and the x and y DACs. The logic could be conveniently located in a DEC BB-11 systems unit inside the computer chassis, connected by cables to the scope. The DMA circuitry consists of an address selector, bus transceiver and multiplexer, interrupt control module, bus master control circuit, and word count and bus address registers. The controller uses the ten device registers listed in Table 5. In order for the processor to communicate with the controller, the controller must be able to respond to the addresses and route information to the appropriate register. The address selector performs this task. Data is received and transmitted from the Unibus using bus tranceivers. The data is multiplexed so that only a single set of transceivers are needed for all ten device registers. This prevents excessive loading of the Unibus by the controller. The interrupt control circuit contains two bus requests modules and a vector address circuit. It is used in two different ways. In order to perform a DMA operation, the Controller must become bus master. One of the request circuits is used to do this. The other is used when interrupts are desired. For example, the operator may wish to change the buffer being displayed. This would best be done at the end of a refresh cycle. An interrupt could be used to inform the processor that the cycle were finished. The vector address circuit is used when generating interrupts. The bus master controller is the circuit that runs the DMA process. When the display controller issues a request for the next data point, the bus master controller directs

Table 5. Register Assignments for Proposed Scope Interface

Name	Description	Address
YSC	Y-Axis Scaling Register	167520
YBUF	Y-Axis Buffer	167522
YSF	Y-Axis Scale Factor	167524
--	Not Used	167526
XSC	X-Axis Scaling Register	167530
XBUF	X-Axis Buffer	167532
XSF	X-Axis Scale Factor	167534
XINC	X-Axis Increment	167536
CSR	Control and Status Register	167540
WCR	Word Count Register	167542
BAR	Bus Address Register	167544

0 0 0 4 3 0 8 0 9 6

the sequence of events. After becoming bus master, a read operation is performed from memory to the controller. The memory location to be read is stored in the bus address register (which is incremented after the transfer to point to the next memory location). The word count register keeps track of the number of points left to be refreshed in that particular refresh cycle. The bus control circuitry is similar to existing DEC hardware.

The controller logic is used to provide timing and signal routing for the display registers. It routes data from memory to the appropriate device register, controls display blanking and unblanking, erases the scope when required and controls the various modes of operation. The user communicates with the controller logic through the control and status register, Fig. 30. If the mode bit (bit 5) is zero, the data set is displayed only once. When completed, the done bit (bit7) is set and if interrupt enable (bit 6) is set, an interrupt occurs. If mode = 1, the data set is displayed, the word count and bus address registers are reset to their original values, and the data set is displayed again. This automatic refresh cycle feature makes programming the controller very simple. In order to display a data set, the bus address register is loaded with the address of the first word of the data set, the word count register is loaded with the number of points to be displayed, other registers may be loaded if required, and the intersity (bit 9) mode, and run (bit 0) bits of the control and status register are set. The data set will then be displayed and refreshed without further

Control and Status Register

15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
Err	Time out	Bus Addr Ovfl	Erase Busy	Single Point	Store	Inter	Erase	Done	Intr Enb	Mode	Toggle	y-Axis Scale	x-Axis Scale	x-Axis Incr	Run

Fig. 30. Bit Assignments for the Control and Status Register of the Proposed Scope Interface.

0 0 0 0 4 3 0 8 0 9 7

intervention. The circuitry required for automatic repetitive scans is based on similar circuitry used in the KW11-P real-time clock.

The control and status register indicates other features of the control logic. The x and y axis circuitry can be run in various modes as indicated below. The control circuitry provides necessary timing and control for the operational mode selected. The scope can be erased by writing bit 9 of the CSR. Setting bit 10 does two things. The scope is put into storage mode. Also the control timing clock is routed through a divider circuit, so that points are displayed more slowly. This is necessary because it takes a longer time to store a point on the screen than it does to illuminate the screen in non-storage mode. For maintainance purposes, a single point mode is available. Writing bit 11 causes a single point to be displayed. Error flags are provided in case of malfunction. These flags are useful because certain types of errors could cause the system to hang up in an intermediate state, halting the processor. Overrides are provided in case an error occurs. Instead of crashing the system an error interrupt is provided.

X and Y Axis Display Registers

We wish to provide two general types of displays in the system, full x-y displays and graph-plot displays. In an x-y display, both the x and y axes can be loaded with any possible data value. In graph-plot mode, we assume that we are plotting a sequence of y-axis values. These values are uniformly spaced along the x-axis by a distance Δx. Graph-plot mode will be the most commonly used mode of operation and we and we therefore will implement the x-axis incrementation in hardware. The x-axis circuit is shown in Fig. 31. In x-y mode

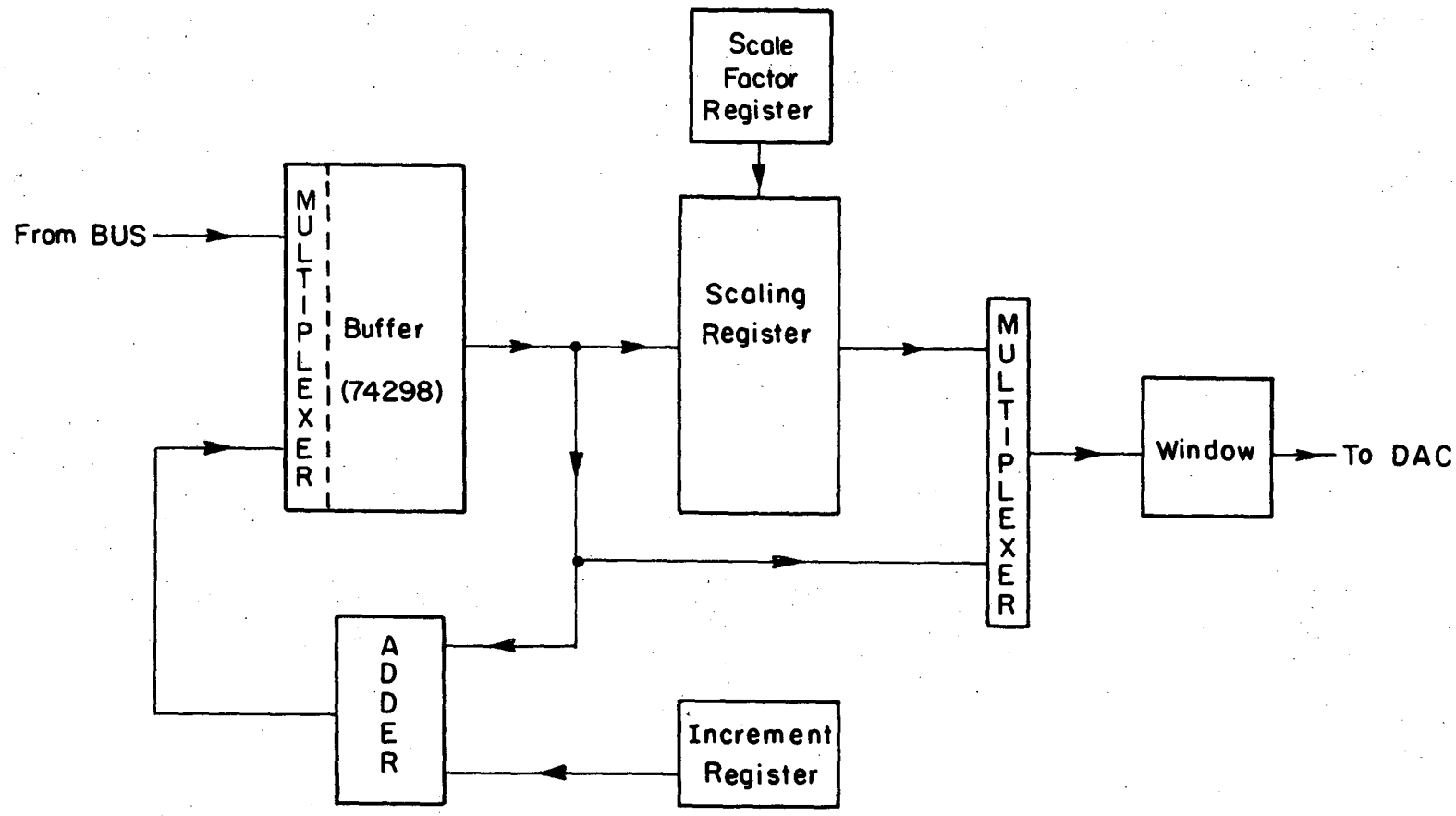


Fig. 31. x-Axis Circuitry--Block Diagram.

0 0 0 0 4 3 0 8 0 9 8

data is loaded into the buffer from the Unibus. The buffer contents may be scaled in the scaling register if the user wishes. The output then is windowed and applied to the x-axis DAC. In graph-plot mode, Δx is loaded into the increment register. The x coordinate of each point is obtained by adding Δx to the present contents of the buffer. The sum is fed back to the input of the buffer to form the coordinates of the new point. Note that the buffer must be edge-triggered in order to prevent the feedback from causing a race condition. The y-axis circuit is identical to the x-axis circuit except that the incremental circuitry is not included.

The scaling register implements the hardware scaling algorithm. The scale factor register contains the offset and gain information required to scale the data. This register is not loaded from the computer. Instead, a series of switches are provided on the panel next to the scope. These switches determine the gain and offset. Suppose we wish to scale the number N to produce N'. Then

$$N' = G\,N + 0,$$

where G is the gain and 0 is the offset. The gain is just $G = 2^m$ where the signed integer m is the quantity input from the switches. The scaling register operation is very simple. The number N is loaded into a bidirectional shift register and m is loaded into a counter. The shift register is then given m clock pulses (at 10-20 MHz) which cause the data to be shifted left ($m > 0$) or right ($m < 0$) by m places. Since a shift is equilvalent to a multiplication or division by 2 in binary arithmetic, this multiplies N by G. The result of the shift

operations and 0 are then added to obtain the scaled result.

The window circuit is used because even after scaling, N' may not be in the range which can be displayed on the scope. Since N' is a 16-bit number, it may take any of 65536 values. The DACs used in the controller are only capable of resolving the 4096 values in the range for 0 to 4095. The window circuit insures that if N' > 4095, the number 4095 is applied to the DAC and if N' < 0, the number 0 is applied to the DAC. This means that out-of-range values will be displayed on the edges of the screen, thereby eliminating possible sources of confusion.

In this appendix I have outlined the design philosophy and circuitry of a display controller which, if implemented, would significantly improve system performance. It would also be very easy to program. Suppose we wished to display n points of a buffer beginning at location b on the screen in refresh mode with operator scaling. We wish to spread the points out across the screen. Therefore the horizontal increment Δx is given by the integer part of 4096/n. The program listed in Table 6 will initiate the display. No further intervention is required until the user wishes to halt the display.

0 0 0 0 4 3 0 8 0 6 6

Table 6. Display Program Example

```
DISPLAY:  MOV     #4096.,MQ ; MQ IS AN EAE REGISTER
          MOV     #n,DIV    ; DIV IS ALSO
          MOV     MQ,XINC   ; QUOTIENT TO XINC
          MOV     #n,WCR    ; NUMBER OF POINTS
          MOV     #b,BAR    ; ADDRESS OF FIRST POINT
          MOV     #1053,CSR ; START DISPLAY.
;RUN, X INCREMENT MODE, Y SCALING, REPEATED REFRESH, INTENSIFY.
```

APPENDIX C.

Command Vocabulary

In this appendix the commands implemented in FTS-11 are briefly discussed. Each command consists of one or two letters and up to four numbers. The letters define the command--they tell the computer what to do. The numbers are parameters necessary for the performance of a given command. For example if we wish to move data from one buffer to another we would have to specify the number of points to be moved. We would also have to specify the source and destination buffers. In the brief command summary that follows these parameters are not listed. Only the function of each command is specified. For more detailed information the reader is referred to the "FTS-11 Operations Manual".

A. Apodize

A	Performs a linear apodization and Fourier transform. The letter A may be followed by the letter of a display command (H, P, T, V).

B. Buffer Editing

B	Lists value of a particular data point on teletype.
BA	Replaces a data point with the average of the previous and following points.
BC	Replaces a data point with a particular value.
BL	Adds write-protection to a buffer. Prevents the data in the buffer from being changed.

0 0 0 0 4 3 0 8 1 0 0

BR Reverses the data points of a buffer. Used if a data set were acquired backwards.

BU Removes write-protection from a buffer.

C. Transform

C Subtracts constant baseline from data set and computes Fourier transform starting with the zero point. Real part of result may be displayed.

D. Deposit Octal Patch

D Used to enter a program modification.

E. Examine Program Location

E Prints the value of a program location. Mainly used for debugging a program.

F. Phase Correction

F Phase correct a data set and transform. Result may be displayed.

FA Compute and display the phase angle of the phase function of a data set. Used to decide if phase correction is required.

FL Phase correct and look. Used when multiple phase corrections may be desired.

G. Coadding

G Coadd. Designed for interferograms. This procedure lines up the maximum points of data sets so that interferograms with different numbers of points before zero path may be averaged.

GS Coadd without the line-up procedure. Used mainly for spectra.

H. Papertape Input and Output

HI Read data from papertape and write it in a buffer.

HM Write data from a buffer to papertape. Place an identifying mark at the beginning of the data set.

HO Write data from a buffer to papertape. No mark is punched.

I. Number of Points

I Prints current number of points in a buffer.

IC Changes point counter indicating number of points in a buffer.

J. Transform

J Takes Fourier transform of a data set without modification.

JM Takes amplitude transform of a data set.

K. Clear Buffer

K Replaces the contents of a buffer with zeros. Sets buffer point counter to zero.

L. Absorption Coefficient

L Replaces the contents of a buffer with the negative of the natural logarithm of the data. The data is assumed to be the result of a Q command.

M. Move Data

M Move a data set from one buffer to another.

0 0 0 0 4 3 0 8 1 0 1

MP Move part of a data set from one place to another.

N. Values of Certain Parameters

NR If acquiring multiple data sets, prints the number of current run being taken.

NV Decode scaling register setting to print gain and offset for the current display.

O. Offset

O Subtracts a constant baseline from a data set. The constant is the average of all the entires in the data set.

OA Performs a linear apodization of a data set.

OL Subtracts a linear drift from the data set. The average of the first and second halves of the data set determine the slope of the line to be subtracted.

OM Perform a "Mertz Apodization" of a data set.

P. Punch Data on Cards

P Punch a data set on cards.

PT Punch a title card.

Q. Ratio of Two Buffers

Q Form the ratio of two data sets. Provision is made for correct numerical values if the data sets were recorded with different gains. The first point of the quotient data set is artifically set to 0.5 to make scaling easy.

R. Take Data

R Acquire a single data set.

RE Extend a run so that a different number of points will be acquired than was called for originally.

RI Collect multiple data sets. After each run, a data tape is punched, the data set is phase corrected and coadded.

S. Not Used

T. Print Data on the Teletype

T Prints the decimal values of a data set on the teletype.

U. Updating Transform

U This command is used to watch the spectrum change as new data points are acquired. The input data is repetitively transformed and the resultant spectrum displayed in storage mode. The result is a family of spectra which display the changing resolution of the data as it is acquired.

V. Visual Display

V Display a data set on the scope.

VS Display a data set on the scope in storage mode. This command is useful, for example, for comparing two data sets.

VE Erase the scope.

VP Plot a data set on the x-y recorder.

W. Write Commands

W Enables comments to be written on the teletype without being interpreted as commands. Used to document a series of operations.

0 0 0 0 4 3 0 8 1 0 2

X. Stop

XC	Stop coadding. Forces the current run to be the last of a series to be acquired.
CNTL/P	Stops display commands such as T, HI, HO and P.
CNTL/X	Aborts data acquisition.

Y. Not Used

Z. Not Used

APPENDIX D.

Data Analysis Algorithms

In this appendix, several algorithms will be presented which have been successfully used in the FTS-11 software for data analysis. In order to understand how these algorithms are coded, we must briefly discuss the way numbers are treated by the software. Numbers are stored in computer memory in the form of one-word signed integers (I format), but this form is often not useful for computation. For example one could not store the cosine of an angle in this format. For some internal operations, a floating point (FP) format is used. This is a two-word format with a one word mantissa and a one-word exponent. For occasions when a two word format is undesirable, two one-word formats are available. In the W format, the single-word integer represents (through a one-to-one mapping) a number in the range [-1,+1). This is the natural format for the Fourier transform algorithm, for example, because it is very convenient to express cosines in this format. For quantities such as ratios, however, which may have magnitudes greater than one, W format is not adequate. We have therefore developed R format. This format is a mapping of a single-word integer into [-8,8).

Fast Fourier Transform

The fast Fourier transform (FFT) algorithm takes advantage of the identities $e^{a+b}=e^a e^b$ and $e^{2\pi i} = 1$ to greatly reduce the number of operations required to perform a Fourier transformation. While many forms of the algorithm exist, we will discuss the "decimation in frequency" version, first suggested by Gentleman and Sande.[1] The coding of this algorithm was first performed by Lee Windmiller[2] and later modified for FTS-11. We consider the calculation of

0 0 0 0 4 3 0 8 1 0 3

$$X(j) = \sum_{n=0}^{N-1} A(n)e(nj/N), \tag{104}$$

where

$$e(nj/N) = \exp(2\pi i\ nj/N). \tag{105}$$

We consider a radix-2 algorithm in which the indices j and n are written as sums of products of powers of two:

$$j \equiv j_{m-1}\ 2^{m-1} + j_{m-2}\ 2^{m-2} + \cdots + j_o \tag{106a}$$

$$\equiv (j_{m-1}, j_{m-2}, \cdots, j_o)\ ,$$

and

$$n \equiv n_{m-1}\ 2^{m-1} + n_{m-2}\ 2^{m-2} + \cdots + n_o$$

$$\equiv (n_{m-1}, n_{m-2}, \cdots, n_o). \tag{106b}$$

Here we have assumed that

$$N = 2^m. \tag{107}$$

Notice that the j_L and n_L are either 0 or 1. In order to keep the subscripts manageable we will specifically calculate an eight point transform. The results can be easily extended. In this case m=3 and $j = j_2 \cdot 4 + j_1 \cdot 2 + j_o$, $n = n_2 \cdot 4 + n_1 \cdot 2 + n_o$.

The sum (104) can be written

$$X(j_2, j_1, j_o) = \sum_{n_o=0}^{1} \sum_{n_1=0}^{1} \sum_{n_2=0}^{1} A(n)e(nj/8). \tag{108}$$

We may factor e in the following way:

$$e(nj/8) = e[(4j_2+2j_1+j_o)(4n_2+2n_1+n_o)/8]$$
$$= e\left(\frac{j_2 n_o}{2}\right) \quad e\left(\frac{j_1(n_o+2_{n1})}{4}\right) e\left(\frac{j_o(n_o+2n_1+4n_2)}{8}\right), \tag{108}$$

and write

$$X(j_o+2j_1+4j_2) = \sum_{n_o=0}^{1} e\left(\frac{j_2 n_o}{2}\right) e\left(\frac{j_1 n_o}{4}\right) .$$
$$\sum_{n_1=0}^{1} e\left(\frac{j_1 n_1}{2}\right) e\left(\frac{j_o(n_o+2n_1)}{8}\right) \sum_{n_2=0}^{1} e\left(\frac{j_o n_2}{2}\right) A(n_o+2n_1+4n_2). \tag{109}$$

In this form the problem becomes one of computing shifted two-point Fourier transforms. The two-point transforms require two complex additions. The shifts each require a single complex multiplication. Let us count operations. If we call a complex multiplication plus a complex addition one operation, then two operations are required to perform a two-point transform and multiply by a phase factor. The algorithm is conveniently divided into three stages or iterations, each of which involves four two-point transforms or eight operations. Therefore 24 operations are required in all. Moreover, the third iteration is particularly simple because the phase factor is unity. This may be compared with direct evaluation of (104) which requires 64 operations.

A few complications arise in a real algorithm. First, the requirement that no "scratch-pad" or extra memory be used for storage of

0 0 0 0 4 3 0 8 1 0 4

intermediate results adds a complication. Because of in-place storage requirements, the result of a transform is stored in bit reversed order. Pairwise interchanges of data points must be performed after the transform is completed. Second, real data may be conveniently treated in this algorithm by clever use of the linearity of the Fourier transform. The procedure for real data is described by Cooley, Lewis, and Welch[3] as well as by other authors. Alternately, Bergland[4,5] has pointed out that a special version of this algorithm could be tailored specifically to real data with high efficiency. This has not been implemented in FTS-11. Finally the algorithm stores the complex result of the algorithm with real and imaginary parts of each data point in sequential memory locations. Since in Fourier spectroscopy we are primarily interested in the real part of the transform, the real and imaginary parts must be sorted out and stored in two separate areas. The final algorithm, including these complications, is capable of transforming a 1024 point data set in about one second using hardware multiplication.

Phase Correction

The problem of phase correction has been discussed by many authors. We will discuss the convolutional phase correction scheme first suggested by Forman and Vanasse.[6,7] For an ideal interferometer, the interferogram is given by

$$F(x) = 2 \int_0^\infty B(\sigma)\cos 2\pi\sigma x \, d\sigma, \qquad (110)$$

where σ is the wavenumber and we have ignored for simplicity effects of finite path length and sampling. If there is dispersion in the interferometer, if (due to sampling) an error is made in the determination

of the zero path point (x=0), or if certain other errors exist, the interferogram will display phase error. That is, the measured interferogram will be asymmetric, given by

$$F_m(x) = 2 \int_0^\infty B(\sigma) \cos[2\pi\sigma x + \phi(\sigma)] d\sigma$$

$$= \int_{-\infty}^{\infty} B(\sigma)\, e^{i\phi(\sigma)}\, e^{i2\pi\sigma x}\, d\sigma. \tag{111}$$

The problem of phase correction is the removal of the effect of $\phi(\sigma)$.

Suppose that we know $e^{i\phi(\sigma)}$. Then the correction of the interferogram is easily performed. We define a convolution function C(x) by

$$C(k) = \int_{-\infty}^{\infty} e^{-i\phi(\sigma)}\, e^{2\pi i\sigma x}\, d\sigma\ . \tag{112}$$

The corrected interferogram may be recovered by performing the convolution

$$F_c(x) = \int_{-\infty}^{\infty} F_m(x')C(x-x')dx'\ . \tag{113}$$

The corrected interferogram $F_c(x)$ can be shown to approximate F(x) by considering the convolution theorem for Fourier transforms. Let us define

$$F * G \equiv \int_{-\infty}^{\infty} F(x')\, G(x-x')dx' \tag{114}$$

and denote the Fourier transform of F by $\tilde{F}$. Then the convolution theorem states that

$$(F * G)^{\sim} = \tilde{F} \cdot \tilde{G}. \tag{115}$$

0 0 0 0 4 3 0 8 1 0 5

Applying the theorem,

$$\tilde{F}_c = (F_m \tilde{*} C) = \tilde{F}_m \cdot \tilde{C}. \tag{116}$$

But

$$\tilde{F}_m = B(\sigma) \cdot e^{i\phi(\sigma)} \tag{117a}$$

and

$$\tilde{C} = e^{-i\phi(\sigma)} \tag{117b}$$

Therefore

$$\tilde{F}_c = B(\sigma)\ e^{i\phi(\sigma)} \cdot e^{-i\phi(\sigma)}. \tag{118}$$

Since F_c has the same Fourier transform as F, a calculation of F_c corrects for phase error that may be present.

Since the calculation of the inverse Fourier transform of $e^{-i\phi}$ and the convolution are easily coded, the problem is reduced to a calculation of $e^{i\phi}$. If ϕ is a slowly varying function of wavenumber, we may calculate $e^{i\phi}$ at low resolution. Suppose the interferogram F_m is measured from $-x_S$ to x_L, where $x_L \gg x_S$. Then we may calculate

$$\hat{B}(\sigma) = \int_{-x_S}^{x_S} F(x) \left[1 - \frac{|x|}{x_S}\right] e^{-2\pi i\sigma x}\ dx \tag{119}$$

and approximate $e^{i\phi}$ by

$$e^{i\phi} \approx \frac{\hat{B}(\sigma)}{|\hat{B}(\sigma)|}\ . \tag{120}$$

Care must be exercised in this calculation, since the process described in (120) is nonlinear. In regions where $\hat{B}(\sigma)$ is small, noise effects can dominate the determination of $e^{i\phi}$. In order to prevent problems with noise, two precautions are necessary. First the resolution used to calculate $\hat{B}(\sigma)$ should be as low as possible for the given $\phi(\sigma)$. We do not wish to resolve $\hat{B}(\sigma)$ very well in regions where $B(\sigma)$ is small. Secondly, $C(x)$ should be apodized before being used in (113). This helps reduce the effect of errors in the calculation of $e^{i\phi}$.

The time required to perform this algorithm is dominated by the evaluation of the convolution integral. If the interferogram is N points long, with m points before zero path, the time required is proportional to mN. Typical 1024 point interferograms may be corrected in 1-2 sec.

Coadding

The usual way of averaging data is to add all of the data and then divide by n, the number of data points. This technique has two disadvantages in our system. First, the stored quantity grows in magnitude as the number of data are increased. This could lead to overflows in a fixed word length environment. Second, the average is available only after the last datum has been recorded. It is desirable to have a running average always available for inspection. The following algorithm is mathematically equivalent to the usual averaging techniques and eliminates these two objections. Let the data be D and the average of the first n values of D be $\langle D \rangle_n$. Then

$$\langle D \rangle_n = \langle D \rangle_{n-1} + \frac{1}{n} [D - \langle D \rangle_{n-1}]. \qquad (121)$$

0 0 0 4 3 0 8 1 0 6

Using this algorithm we may simply store $\langle D \rangle_n$ and then update it every time a new D is measured. Equation (121) may be evaluated in about 20 μs.

Special Functions

In various computations it is necessary to calculate square roots, natural logarithms, and arc tangents. For example, the calculation of $e^{i\phi}$ in phase correction requires a square root, display of the phase function ϕ requires an arc tangent, and absorption coefficient calculations require a natural logarithm. Standard techniques are used to calculate these functions. The interested reader is referred to the book by Hart.[8]

References

1. W. M. Gentlemen and G. Sande, Fall Joint Computer Conference Proceedings 29, 563 (1966).
2. L. Windmiller, private communication.
3. J. W. Cooley, P. A. W. Lewis, and P. D. Welch, The Fast Fourier Transform Algorithm and Its Applications, IBM Watson Research Paper RC-1743 (1970), pp. 30-33.
4. G. D. Bergland, Commun. ACM 11, 703 (1968).
5. G. D. Bergland, IEEE Transactions AU-17, 138 (1969).
6. M. L. Forman, W. H. Steel, and G. A. Vanasse, J. Opt. Soc. Am. 56, 59 (1966).
7. H. Sakai, G. A. Vanasse, and M. L. Forman, J. Opt. Soc. Am. 58, 84 (1968).
8. J. F. Hart, Computer Approximations (John Wiley and Sons, New York, 1968).

0 0 0 0 4 3 0 8 1 0 7

LEGAL NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Energy Research and Development Administration, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

TECHNICAL INFORMATION DIVISION
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720