A FABRY-PEROT SPECTROMETER

FOR

AURORAL OBSERVATIONS

A Thesis

Submitted to the Faculty of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree of

Master of Science

in the Department of Physics

University of Saskatchewan

by

Written under the Supervision of

Saskatoon, Saskatchewan

October, 1960

ACKNOWLEDGMENTS

The author is grateful to Dr. B.W. Currie and the Physics Department for the facilities made available, and the University of Saskatchewan for the financial assistance in the form of an Assistantship.

The author would like to show his appreciation to Dr. G.G. Shepherd for the guidance and ideas that made this thesis possible.

Without the multilayer evaporation facilities made available by Dr. D.M. Hunten for reflecting coatings on the plates and the fine interference filters, this work could not have been as extensive.

Mr. F. Rittman, of the Physics Department machine shop, is thanked for his fine workmanship, keen interest and good ideas which helped in the design and fabrication of this instrument.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE page

LIST OF TABLES

I INTRODUCTION

Spectrometers utilizing prisms or gratings have been in use for a number of years. Interferometers such as the Fabry-Perot and Michelson are often considered as being extremely delicate and specialized instruments, not the kind of instrument one would choose for routine observation of spectra. Recently, however, special properties of the Fabry-Perot have been recognized and a very useful and versatile spectrometer designed, which puts the Fabry-Perot ahead of other spectrometers in many respects.

The Fabry-Perot was invented in the nineteenth century, and used almost immediatly in wavelength measurements and standardization of the meter. The Fabry-Perot interferometer has also been used a great deal for hyperfine structure studies. It has always been recognized as being a useful instrument because of its properties of being able to convert directly from wavelength of light to a standard length and the fact that its theoretical resolution has no limit. The Fabry-Perot is a multiple beam apparatus that splits the beam up by successive reflections between two plates, at each reflection allowing some of the beams to pass through, thus making many beams interfere. The result is that the interference fringes are very sharp.

The method by which a Fabry-Perot has been and still is often used to measure wavelengths and study fine structure, is what will

be called here, the photographic method. In this method an objective lens focuses the Fabry-Perot fringes, which are at infinity, onto a photographic plate. When looking at a monochromatic source the image consists of a number of concentric circles which become closer together as one proceeds from the center. The Fabry-Perot is often crossed with a prism or grating spectrometer, because the free spectral range is small, when something other than a nearly monochromatic source is being used. These fringe patterns are quite difficult to analyse but for many years it was the best method to get such high resolution, and it is still used.

The photoelectric method of recording Fabry-Perot fringes was developed by Jacquinot and Dufour (1949). The method used is to replace the photographic emulsion at the focus of the objective lens with a diaphragm that allows a portion of the pattern to go through, be collected by a lens and be focused on a photomultiplier tube. If operated in this manner it is possible, by changing the optical path length between the plates, to make the instrument scan its passbands linearly with time over a wavelength, or wavenumber, interval; thus it is comparable to other spectrometers. The recorded output is the spectrum, unlike the photographic method where considerable analysis is required to get a spectrum.

Other advantages of a Fabry-Perot spectrometer compared to a photographic Fabry-Perot are discussed below. A spectrum can be obtained in a shorter period due to higher sensitivity of the

photomultiplier which is especially important in aurora because of the short life of some forms that have interesting spectral characteristics. The developing of plates and subsequent reduction to intensities by microphotometer tracings is eliminated, which is very important. This means the non linearities in emulsions are overcome, the time required to get data is reduced, and there is an increase in accuracy due to direct photoelectric methods. The objective lens does not have to be of high quality since only near-axial rays are used, and a high f/number to get high speed is not required (in fact at high resolution a lens of long focal length is advantageous).

An important advantage that any spectrometer has over a spectrograph is that it is possible to see the results as the observations are being taken, and thus observing time can be used more efficiently. If rapidly moving sources are being studied perhaps a photographic method has some advantages, but if the spectrum is changing rapidly this advantage is lost.

The great advantage in light gathering power for a given resolving power for a Fabry-Perot spectrometer, over prism and grating spectrometers, is clearly pointed out by Jacquinot (1954). He compares instruments of the same effective area and resolving power and finds the grating spectrometer always will have greater light gathering power than a prism spectrometer, and a Fabry-Perot spectrometer will have as much as 30 to 400 times the light gather-

ing power as a grating spectrometer. This shows why the Fabry-Perot can be superior even for low resolution where good light gathering power may be important. A Fabry-Perot has further advantages over a grating spectrograph in that one set of plates can be used for any wavelength region in which they will transmit, and the resolution can be adjusted to any value. This is not true for a grating spectrograph since a particular grating is designed to be most efficient in one spectral region and at one resolution. This is why the Fabry-Perot spectrometer is a very versatile instrument.

In upper atmospheric observations until very recently, the Fabry-Perot has been limited to measuring wavelengths. The classic examples using the photographic method are the exact measurement of wavelength of the auroral green line by Babcock (1923), and the confirmation of the sodium D lines in the twilight by Bernard (1938). Others were the study of both 5577A and 6300A by Vegard (1937) and a study of all of these by Dufay, Cabannes and Gauzit (1942).

Babcock (1923) realized that the Fabry-Perot could be used for temperature measurement even though it was not known at the time what was producing the oxygen 5577A auroral and night air-glow green line. The method of obtaining temperature is called the Doppler method. The profile of the spectral line is a pure Gaussian and results from the Doppler shift arising from the

the atoms random motion in thermal equilibrium. The half-width s, and absolute temperature T are related by $s = 7.16 \times 10^{-7} \sigma \sqrt{\frac{T}{M}}$ where σ is the wavenumber of the line in cm^{-1}, and M is the molecular weight. To give good results when using the Doppler method a line should have narrow natural half-width, due to internal structure and transition probability. To fill the second need a forbidden transition is best because of its long time in the excited state. Another important consequence of the forbidden transition is that the atoms spend a sufficiently long time in the excited state to ensure that they will reach thermal equilibrium. The oxygen 5577A [OI] fills these requirements since it is forbidden, having life time of about 0.7 seconds in the upper state, and is not known to have any internal structure that would broaden the line compared to the Doppler width. Another important requirement is that it be bright enough so that enough luminosity can be obtained to use high resolution. The oxygen green line is one of the most prominent features in the aurora. Babcock assigned an upper limit to the half-width of the green line of 0.035 A. Vegard (1937) tried to get a temperature for the oxygen red line (6300A), but could not obtain enough resolution with the luminosity required.

Because of this there was a time lapse of almost twenty years where the idea of getting temperatures by this method was given up. With the advent of dielectric multilayers, the absorption

in the reflecting layers was greatly reduced and interest was again aroused. Wark and Stone (1955), Wark (1956), (1960), and Cabannes and Dufay (1955), (1956a), (1956b), resumed photographic work on the aurora and night airglow using a Fabry-Perot, and found it ~~is~~ possible to obtain temperatures in this way.

When the Fabry-Perot interferometer was investigated by Jacquinot (1954) interest was aroused in this method and some Doppler temperatures were tried again. The reason for the difficulty experienced by these many observers is the high resolution required. Using the half-width obtained by Babcock (1923), and assuming the instrument would need a passband at least as narrow, the minimum resolution required becomes $\frac{5577A}{0.023A} = 2.4 \times 10^5$. However, with the increase in sensitivity of photomultipliers and of course dielectric layers, it was thought worth while to try to get Doppler temperatures. Armstrong (1956), (1959) used a Fabry-Perot spectrograph to try and measure the Doppler temperature of the 5577A line in night airglow and a bit of aurora and he obtained some preliminary results. Karandikar (1956), (1959) has an instrument capable of measuring Doppler temperatures, but at the time of writing his results are not known since he had not the opportunity to observe aurora until recently.

It is considered that Doppler half-width measurements of the auroral green line, 5577,^{A} would give the most reliable spectroscopic temperatures of those available, since the interpretation

is beyond question. The other methods of obtaining temperatures are vibration and rotation. Vibrational temperatures are not very reliable because, for temperatures below 1000°K the excitation process dominates and largely determines the population of the upper states, rather than temperature. Rotational temperatures give reasonable values but the interpretation depends upon the excitation process. Thus Doppler temperatures should be the best and the oxygen green line should give very reliable results. The reason that more temperatures have not been obtained by the Doppler method is that high resolution is required that cuts down the intensity since the product of luminosity and resolution is a constant for spectrometers.

Because of the obvious suitability for auroral studies, it seemed desirable to build a Fabry-Perot spectrometer at the University of Saskatchewan. Of the possible studies, the most worth while appeared to be a high resolution instrument that would be capable of observing aurora for the oxygen green line with Doppler temperatures in mind and which would be easily convertable to other studies. Other observers have had little chance to observe aurora where the gain in intensity over night airglow is great. A pair of 4 inch diameter quartz plates were ordered from Hilger and Watts before the author undertook this thesis. They arrived in the winter of 1959 and this study was begun in the spring, ~~thus~~ A Fabry-Perot interferometer was to be built with

the first purpose in mind to study the auroral green line for temperatures.

The next chapter consists of some theory required for subsequent chapters but it is not intended to be complete. Next in Chapter III a full description of the apparatus is given. This is described in detail in most parts because the design is thought to be original and appears very satisfactory. In Chapter IV the contours of the plates are determined. This is important because of the great effect the defects of the plates have on the way in which the plates are to be used and on the end result. The next chapter contains some results on auroral observations which are of a preliminary nature, but thought to be worth while. The final chapter ~~is the conclusions that~~ gives an estimate of the value of the results and a few ideas on what may be done next with the four-inch Fabry-Perot interferometer.

II THEORY

2.1 Introduction

In this chapter, enough theory will be given to explain the calculations in the following chapters. The elementary theory may be found in Jenkins and White (1950), but a much more extensive treatment is in Chabbal's (1953) paper. The notation used will be Chabbal's except for a few cases. A definition of symbols used is collected in Table 2.1.

Refering to Figure 2.1 a brief description of the apparatus will be given. The plates are made of high quality quartz carefully polished flat and have partially reflecting coatings on the faces that are adjacent to each other. A single ray from the source B at the angle i to the optical axis will be followed through the interferometer. This ray, some of which passes through the first plate, is reflected between the plates, each time allowing a portion to continue on through the second plate, forming a bundle of rays. These rays which are still all at the same angle i interfere more or less depending upon the phase shift between adjacent rays. These rays because they are all parallel appear to come from infinity, so the objective lens focuses them all at one spot on the focal plane. It is easily seen that since the phase change depends upon the angle i, the fringes will be circular, and concentric about the optical axis. If a photographic method were

TABLE 2.1

A - Airy function: a - half-width of A

$\mathcal{A}$ - absorption in reflecting layers

B - source function: b - half-width of B

D - defect function: d - half-width of D

E - etalon function: e - half-width of E

F - aperture function: f - half-width of F

i - angle of ray from normal of plates

L - luminosity

m - order of interference, m_o if cos i is negligible

mk- millikaiser (1/1000 wavenumber (cm^{-1}))

N - finesse: N_R - reflection, N_D - defect, N_E - etalon

n - index of refraction

R - reflectivity of plates

$\underline{R}$ - resolution

S - surface area of plates

t - distance between plates (spacing)

W - instrumental profile: w - half-width of W

Y - registered function: y - half-width of Y

$\Delta\sigma$- separation between orders

θ- angle subtended by aperture from objective lens

σ- wavenumber of spectral line, σ_i - peak wavenumber

ϕ - phase shift

τ^a- absorption coefficient

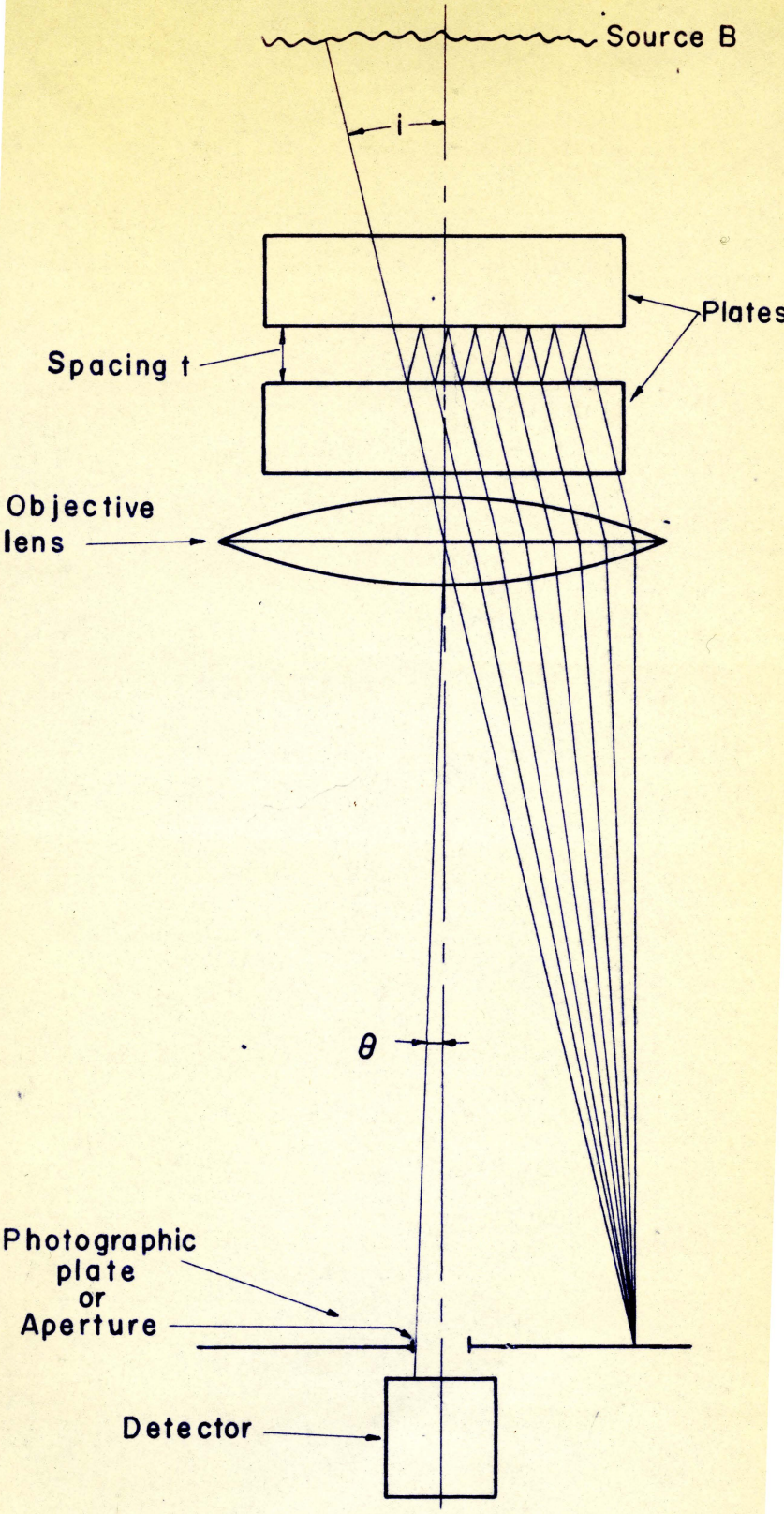

Source B
i
Plates
Spacing t
Objective lens
θ
Photographic plate or Aperture
Detector

FIGURE 2·1

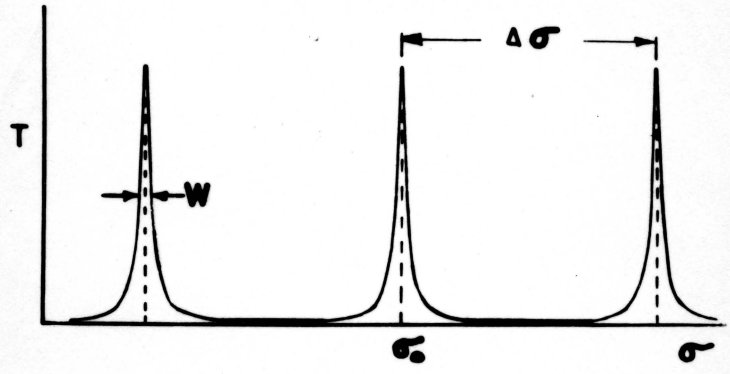
to be used, as has been done for many years, a photographic plate would be put at this focal plane to record the interference fringes.

2.2 Airy Function

The function of transmission of the plates for a given phase change ϕ is derived in Jenkins and White (1950), and is called the Airy function. This is plotted in Figure 2.2, with the transmission as a function of σ, the wavenumber $(\frac{1}{\lambda})$. ϕ can be put in terms of σ as shown below;

$$\text{Airy function} \equiv A = \frac{1}{1 + p \sin^2 \frac{\phi}{2}}$$

where p is a constant, which in the case of the Fabry-Perot is equal to $\frac{4R}{(1-R)^2}$, R being the reflectivity of the plates. ϕ is the phase shift between adjacent beams as seen from Figure 2.1 and is given by $\phi = 4 \pi n t(\cos i)\sigma$. The variables in this formula are the refractive index between the plates n, the distance between the plates t, the angle of incidence i, and the wavenumber σ of the spectral line being studied. The order of interference m, can be defined as being $\frac{\phi}{2\pi}$, the number of complete phase shifts, thus $m = 2 n t(\cos i)\sigma$. If we let σ_i be the wavenumber at the peak of the Airy function then this corresponds to m and we can write $\frac{m}{\sigma_i} = 2 n t \cos i$. If we keep everything on the right of the equation constant, then as we go from σ_i to σ which is a small amount off the peak, m must also change to $m + \epsilon$ where ϵ is a fraction of an order, thus we can say,

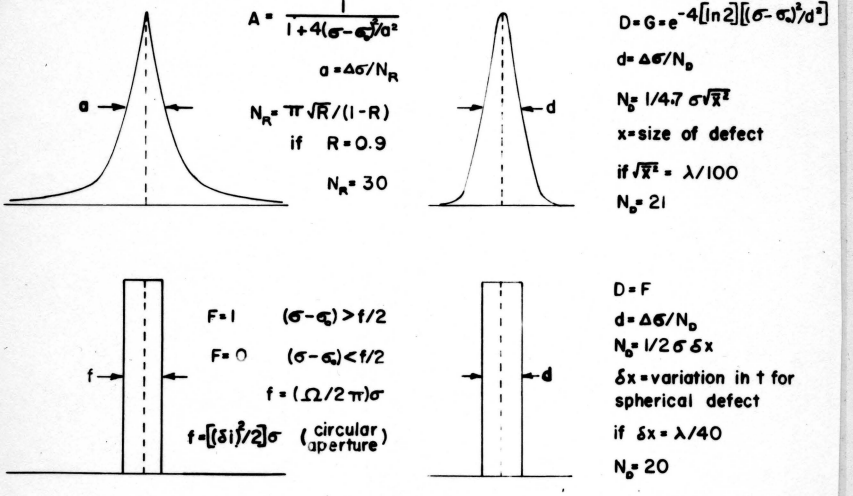


$T = F[t, n, i, (\sigma - \sigma_0)]$

$N = \Delta\sigma / W$ (Drawn with N = 20)

$\Delta\sigma = 1/2\,nt$

Airy Function



FUNCTIONS DETERMINING FABRY-PEROT PASSBAND

FIGURE 2.2

$$\frac{m}{\sigma_i} = \frac{m+\epsilon}{\sigma} \quad \text{and} \quad \epsilon = \frac{m}{\sigma_i}(\sigma - \sigma_i)$$

If this fractional order $m+\epsilon$ corresponds to $\frac{\phi}{2\pi}$ then

$$\phi = 2\,n\,t\cos i\,2\pi\sigma = \frac{m+\epsilon}{\sigma}\,2\pi\sigma = 2\pi m + 2\pi\epsilon$$

so $\sin\frac{\phi}{2} = \sin\pi m\cos\pi\epsilon + \cos\pi m\sin\pi\epsilon = \sin\pi\epsilon$

Substituting for ϵ and squaring gives

$$\sin^2\frac{\phi}{2} = \sin^2\left[\frac{\pi m}{\sigma_i}(\sigma - \sigma_i)\right].$$

Replacing $\frac{m}{\sigma_i}$ by $2\,n\,t\cos i$ and substituting in the formula for the Airy function gives;

$$A = \frac{1}{1 - \frac{4R}{(1-R)^2}\sin^2\left[2\pi n t(\cos i)(\sigma - \sigma_i)\right]}$$

In the case of photoelectric scanning i is usually very small (in our case $< 0.17^\circ$) and the final form of the formula is

$$A = \frac{1}{1 - \frac{4R}{(1-R)^2}\sin^2\left[2\pi n t(\sigma - \sigma_i)\right]}$$

which puts A in terms of σ as was required.

In this thesis, the symbol $\Delta\sigma$ will be reserved for the separation between passbands and is found by letting ϕ increase by 2π and finding the corresponding change in σ. Using

$$\phi = 4\pi\,n\,t\,(\sigma - \sigma_i)$$

and differentiating keeping n and t constant gives

$$\Delta\phi = 4\pi\,n\,t\,\Delta\sigma = 2\pi.$$

Hence $$\Delta\sigma = \frac{1}{2nt}.$$

If the Airy function is set equal to $\frac{1}{2}$ then the width of the function at half-height, called the half-width a, can be found. This is given by $a = \frac{(1-R)\Delta\sigma}{\pi\sqrt{R}} = \frac{\Delta\sigma}{N_R}$, where N_R is a constant depending only on the reflectivity of the plates. It is called

the reflective finesse and is a very important concept since it only depends upon the reflectivity of the plates and not their separation. If there is absorption in the reflecting layers, the Airy function is multiplied by a constant $\tau^a = \left(1 - \frac{a}{1-R}\right)^2$, where a is the absorption coefficient and R the reflectivity. τ^a is like a scaling factor that lowers the peak transmission but leaves the shape of the curve unaltered.

2.3 Scanning

As long as n and t are constant, there is an absolute scale for A plotted against σ . If however, either n or t is varied the whole array of passbands can be made to slide along the σ scale, thus the instrument can be made to scan its passbands over a wavenumber interval. This makes possible a photoelectric interferometric spectrometer since a small portion of the fringe pattern can be allowed to pass through an aperture and strike the cathode of a photomultiplier tube. This means that the output from the amplifier and chart recorder will be a spectrum of the source, modified by the instrument.

Since the Fabry-Perot fringes are circular, obviously a circular diaphragm concentric with the fringes should be used, otherwise light would be lost at no gain in resolution. The simplest circular diaphragm ~~is~~ has a hole, or aperture, at the center of the pattern which includes the angle i from 0 to θ .

If A has maximum transmission at i = 0, then the order m

must be integral and it can be written as $m_o = 2\,n\,t\,\sigma$ where m_o is used since the cosine term can be neglected when i is small. There are then, only two ways to vary σ, by varying n or t. If n is varied, then by differentiating $m_o = 2\,n\,t\,\sigma$ the corresponding change in σ can be found.

$$0 = 2\,d\,n\,t\,\sigma + 2\,n\,t\,\,d\sigma \quad \text{giving} \quad \frac{dn}{n} = -\frac{d\sigma}{\sigma}.$$

This means that $d\,n$ = constant X $d\sigma$ as long as $n >> d\,n$ and $\sigma >> d\sigma$. Hence the wavenumber interval covered depends directly upon the change in refractive index and is independent of the plate spacing.

If t is varied the equation becomes after differentiating

$$0 = 2\,n\,d\,t\,\sigma + 2\,n\,t\,\,d\sigma$$

giving

$$\frac{dt}{t} = -\frac{d\sigma}{\sigma}$$

For a given spacing $d\,t$ = constant X $d\sigma$ as long as $t >> d\,t$ and $\sigma >> d\,\sigma$. However, the spacing is important in this case because for a given d t the wavenumber scan is inversely proportional to the spacing. Another interesting way to look at this difference in the two methods is to let the order m_o change (an integral number of times if desired) with constant σ, and hence see how many passbands can be covered with a given change in n or t. Differentiating $m_o = 2\,n\,t\,\sigma$ first with respect to n we have $dm_o = 2\,dn\,t\,\sigma$ which shows that the number of orders gone through is directly proportional to the spacing, as it of course must be since the wavenumber interval is independent of the

spacing. Then differentiating with respect to t we have, $dm_o = 2\ n\ dt\ \sigma$ since n and σ are constant, this shows that the number of orders that one can go through when scanning by this method depends directly upon dt, and is independent of the spacing.

Thus at low order when t is small, to scan over a large wavelength interval changing t by mechanical scanning would be the best, since $\Delta\sigma$ corresponding to a change of one order is large. At high order the refractive index scanning method appears the best to scan over a large wavelength interval. It is possible with the apparatus described in the next section using Freon-12 as the gas for refractive index scanning to scan over about 60 orders with the spacing t = 1 cm. It has been found possible only to scan over a few orders using mechanical scanning (by varying t) as the plates can not be kept in adjustment when moved very far. However at low order, if t = 0.17 mm, then using the same refractive index change as above it would be possible to scan over only one order and mechanical scanning would be preferable. There is thus a point below which mechanical scanning is best and above which refractive index scanning is the best. This point is around t = 1 mm. Since the refractive index method is technically the easiest especially for plates of large area it may be worth while to use refractive index scanning even when the spacing is small, by making the instrument capable of withstanding high pressure changes.

2.4 Defects

What is required as the end result is the spectrum of the source. If the instrument had a passband infinitely fine, of course the output of the machine would be the same as the input. But even if this were possible, no light could get through. The instrument tends to broaden the source function B and we can call the broadening function, the instrumental profile W. Up to now we have assumed that the Airy function was the only function making up the instrumental profile and in this case $W \equiv A$. However, there are other factors and the next one to consider is the defects of the plates.

The defects are caused by curvatures in the plates, roughness left after polishing, and the fact that the plates may not be in adjustment. An analytical function representing the defects in the plates can be set up as follows. It is assumed that the plates are made up of a large number of infinitesimal etalons that have a spacing t' that varies above and below the mean value t. A value of σ, denoted as σ', can be associated with this t' by using the formula $m_o = 2 n \sigma t$. This σ' is slightly different from the σ denoted by the pure Airy function, depending upon how much t' varies from the t which was assumed for the perfect plates. The defects then broaden the Airy function since for each value of σ in the perfect case there is a range of values of σ' that transmit in proportion to the area of the plates at the spacing t'.

We call $D(\sigma')$ the defect function and make it equal to the area of the portion of the plates that has spacing from t' to $t' + dt'$ which is associated with σ' to $\sigma' + d\sigma'$. Thus for any value of σ', $D(\sigma')$ gives the element of area associated with an element of wavenumber $d\sigma'$ around σ'. If the defect function is summed over all σ' it will of course give the total area of the plates S.

$$S = \int_{\sigma'=0}^{\infty} D(\sigma')\, d\sigma'.$$

To obtain a function for the imperfect Fabry-Perot plates, the Airy function at the wavenumber σ' is multiplied by $D(\sigma')$ which gives the effective area for that wavenumber σ', summed over all σ' and divided by the total area of the plates to normalize it again. This function will be called the Etalon function and can be written

$$E(\sigma - \sigma_i) = \frac{1}{S} \int_{\sigma'=0}^{\infty} A\left[(\sigma - \sigma_i) - \sigma'\right] D(\sigma')\, d\sigma'.$$

An integral of this type is the convolution of two functions and may be written as $A \ast D$. Thus the etalon function E is the convolution of A and D divided by the total area of the plates. An important property of this kind of product, like $C = A \ast B$, is that the Fourier transform of C is equal to the product of the transforms of A and B. The halfwidths of these functions, denoted by lower case letters, combine in different manners depending upon the function, and are written as $c = a \oplus b$. That is, the convolution of a function A of half-width a with a function B of half-

width b results in a function C which has half-width c.

A finesse N_D can be assigned to the defect function. It was noted in the paragraph on scanning that a given variation in t results in a certain change in order, independent of the spacing t. Thus a small change in t, t-t' corresponding to the half-width d of the defect function, will always be the same fraction of the change in t required to change the order by one unit. Thus d will always be a given fraction of $\Delta\sigma$ and hence a finesse $N_D = \frac{\Delta\sigma}{d}$ can be defined which is a constant for given plates. It follows that the etalon function also has a finesse N_E defined as $N_E = \frac{\Delta\sigma}{e}$ where e is the half-width of the etalon function. The half-width and finesse tend to lose their meaning if the function being considered is not symmetrical, as is possible for the defect function and etalon function. It is possible to have a defect function not symmetrical (as will be shown in Chapter IV) but since the value for t is the mean value of t' (weighted according to area), D tends in general to be nearly symmetrical.

2.5 Detector Function

The instrumental passband, represented so far by the etalon function E, is further broadened because the aperture has a finite size. This effect can be represented by the detector function F, that depends upon the angle θ subtended by the aperture from the objective lens. This function, unlike the Airy function and defect function, has width f independent of plate spacing t.

Because of this a true finesse can not be assigned to it, but it is sometimes useful to think of a finesse $N_f = \frac{\Delta\sigma}{f}$ which is valid only for a given plate spacing, since most work is done with a constant spacing t. F thus depends upon the focal length of the objective lens and the diameter of the aperture.

The transmission of F is then a rectangular function when plotted against i since it transmits 100% from i = 0 to $i = \theta$ and zero for $i > \theta$. The width of f can be related to σ in the following manner. $f = \delta\sigma$ (a wavenumber interval) so if $\delta\sigma$ can be put in terms of θ then F can be used as a function of σ like the Airy function and defect function.

Differentiate $m_o = 2\,n\,t\,\sigma \cos i$, keeping m, n, and t constant.

$$0 = 2\,n\,t\,d\sigma \cos i - 2\,n\,t\,\sigma \sin i\, di$$

giving $\frac{d\sigma}{\sigma} = \frac{\sin i}{\cos i}\, di$.

Now using approximations cos i = 1, sin i = i and σ = constant (which are all very good in the case of high resolution such as the instrument being regarded here where $\theta < 0.17^\circ$, and σ is 17,900 cm^{-1} compared to $\delta\sigma < 0.1\ cm^{-1}$), integrate.

$$\frac{1}{\sigma}\int_{\sigma}^{\sigma+\delta\sigma} d\sigma = \int_0^{\theta} i\, di$$

giving $\frac{\delta\sigma}{\sigma} = \frac{\theta^2}{2}$ hence $f = \frac{\sigma\,\theta^2}{2}$

or in terms of the solid angle Ω subtended by the aperture from the objective lens $f = \frac{\Omega\,\sigma}{2\pi}$.

The convolution of F with E can be performed giving $W = E \ast F$. W is called the instrumental profile with w its half-width. It is

this function that operates on B to give Y, the recorded function. This is again a convolution, represented by $B \not\ast W = Y$. This can all be broken down again in reverse order as shown by $Y = B \not\ast W = B \not\ast E \not\ast F = B \not\ast A \not\ast D \not\ast F$. These products are commutative and associative, hence $B \not\ast (E \not\ast F) = B \not\ast E \not\ast F = F \not\ast E \not\ast B$ etc. The half-widths can also be broken down in a similar manner.

$y = b \oplus w = b \oplus e \oplus f = b \oplus a \oplus d \oplus f = f \oplus d \oplus b \oplus a$.

2.6 Analysis of Recorded Function Y

In order to find B from Y, the function W must be known. If the functions A and D (or E) and F are known then their Fourier transforms can be found and thus the Fourier transform of W will be known. ~~This~~ W is constant except for the fact that D changes with temperature, and most important, with adjustment. W can also be measured experimentally by observing a known spectrum B', before and after the unknown spectrum.

In some cases the half-width of the source is all that is required, for instance in temperature measurements where it is assumed that the source profile B is pure Gaussian. If b is known, then by using $b = 7.16 \times 10^{-7} \sigma \sqrt{\frac{T}{M}}$, T may be found. The manner in which the half-widths combine is given by Chabbal (1953), but some are quite simple and will be discussed here. If two Airy functions are combined that have half-widths a_1 and a_2, the result $a_3 = a_1 \oplus a_2$ is given simply by $a_3 = a_1 + a_2$. If two Gaussian (G) functions of half-widths g_1 and g_2 combine, the

resulting half-width is given by $g_3^2 = g_1^2 + g_2^2$. If two rectangular functions with widths f_1 and f_2 combine the result is a triangular function if $f_1 = f_2$ with half-width $f_3 = f_1 = f_2$, and if $f_1 > f_2$ the result is a trapezoidal function with half-width $f_3 = f_1$. Thus it can be seen that the Airy function tends to broaden a given function the most, the Gaussian function next and the rectangular function the least. The other cases such as A ✗ G, A ✗ F, G ✗ A, G ✗ F are dealt with in Chabbal's paper and graphs are drawn that enable one to find the third half-width in $c = a \textcircled{+} b$ when two are known.

2.7 Luminosity and Resolution

The luminosity L is defined as the maximum light flux received by the photomultiplier through the aperture, assuming the instrument is analysing a single spectral line. It is thus the height of the recorded function Y. The resolution is defined as $\frac{\sigma}{y} = R$ when the interferometer is analysing a line with half-width infinitely narrow.

The luminosity can be written $L = B\ S\ \Omega\ \tau$ where B is the source brightness, S the surface area of the plates and Ω the solid angle accepted by the plates, τ being the transmission at the peak of the passband. From Section 2.5 it was found that $f = \frac{\Omega \sigma}{2\pi}$. Substituting for Ω in the formula for L gives

$$L = B\ S\ 2\pi \frac{f}{\sigma} \tau$$

The resolving power is roughly equal to $\frac{\sigma}{f}$, although it will be

lower of course, but this difference can be absorbed into τ. This then gives $L = \frac{B52\pi\tau}{R}$ and hence for a given source L X R is constant.

It is found that with given plates and reflectivity so that N_R and N_D are fixed τ can be maximized by picking the proper f, which seems reasonable since it would be wasteful to have f much larger or smaller than e. Hence, to keep f the proper value as the plate spacing is changed, to change the resolution, the finesse $N_F = \frac{\Delta\sigma}{f}$ is kept constant (by varying f) and nearly the same value as N_E, which stays constant since it is a true finesse. To increase the resolution, then, what must be done is move the plates farther apart keeping the same finesses, and then R will increase at the expense of L. With different N_R and N_D (obtained by different reflection or different areas of the plates utilized) and the proper N_F, τ will have changed so that L X R will again be a constant for this set of values for the finesses, but the constant may be larger than before. For example it may be possible to black out a part of the plates that is contributing a lot to the widening of the defect function such that the increase in resolution will more than make up for the loss in luminosity.

As a general rule the Airy function should be slightly narrower than the defect function since it broadens the line more, and the width f of the detector should be made such that $f \doteq e$, the half-width of the etalon function. In a particular case this

depends upon the problem being studied and the actual shape of the defect function. It is also advisable to work with the plates as far apart as possible to get the required resolution without danger of overlapping of orders. In this case N_R and N_D don't have to be very large, and the quality of the plates is not as important. If however, high resolution is required and the source being studied has complicated structure, the plates can't be too far apart or $\Delta\sigma$ will be smaller than the structure being analysed, and then to get the required resolution N_R and N_D must be high. The instrument will not give as high a value for L X $\underline{R}$ as is the case when the plates are farther apart and N_F decreased. This can be overcome if the Fabry-Perot is preceded by a monochromator that is capable of cutting out the other passbands that are not wanted. This may be in the form of a prism, or grating spectrograph or another Fabry-Perot. It is interesting to note that if either a grating spectrograph or Fabry-Perot is put in the same chamber as the main Fabry-Perot, they can all be scanned at the same time by using refractive index scanning, as once they are adjusted no complicated scanning mechanism is required to make both instruments scan at the same rate. See Hershberg and Kadesh (1958).

III DESCRIPTION OF APPARATUS

3.1 Introduction

An assembled drawing of the apparatus is given in Figure 3.1. Each component will be described in detail in the following sections; here, only a general description will be given.

Starting on the inside there are the Fabry-Perot plates made partially reflecting on the faces adjacent to each other. The plates are kept separated and parallel to each other by the spacer and adjusting mechanism on the plate holder. The term etalon has been reserved for the complete assembly of the plates, the spacer, the plate holder and the adjusting mechanism. The etalon must be very stable so the plates will stay in adjustment.

The etalon is mounted in the pressure cell which serves as a chamber in which the pressure and hence the index of refraction can be changed. The pressure cell is wound with heating coils so that the whole system inside can be temperature regulated. In order to orient the etalon so that its optical axis is vertical and in line with the axis of the objective lens, the whole pressure cell must be moveable. This is accomplished by having the pressure cell rest on a platform which in turn rests on three screws that go through the base plate. Thus the pressure cell and hence the etalon can be tilted. The horizontal position is determined by three adjusting screws mounted on the platform that touch the side of the pressure cell.

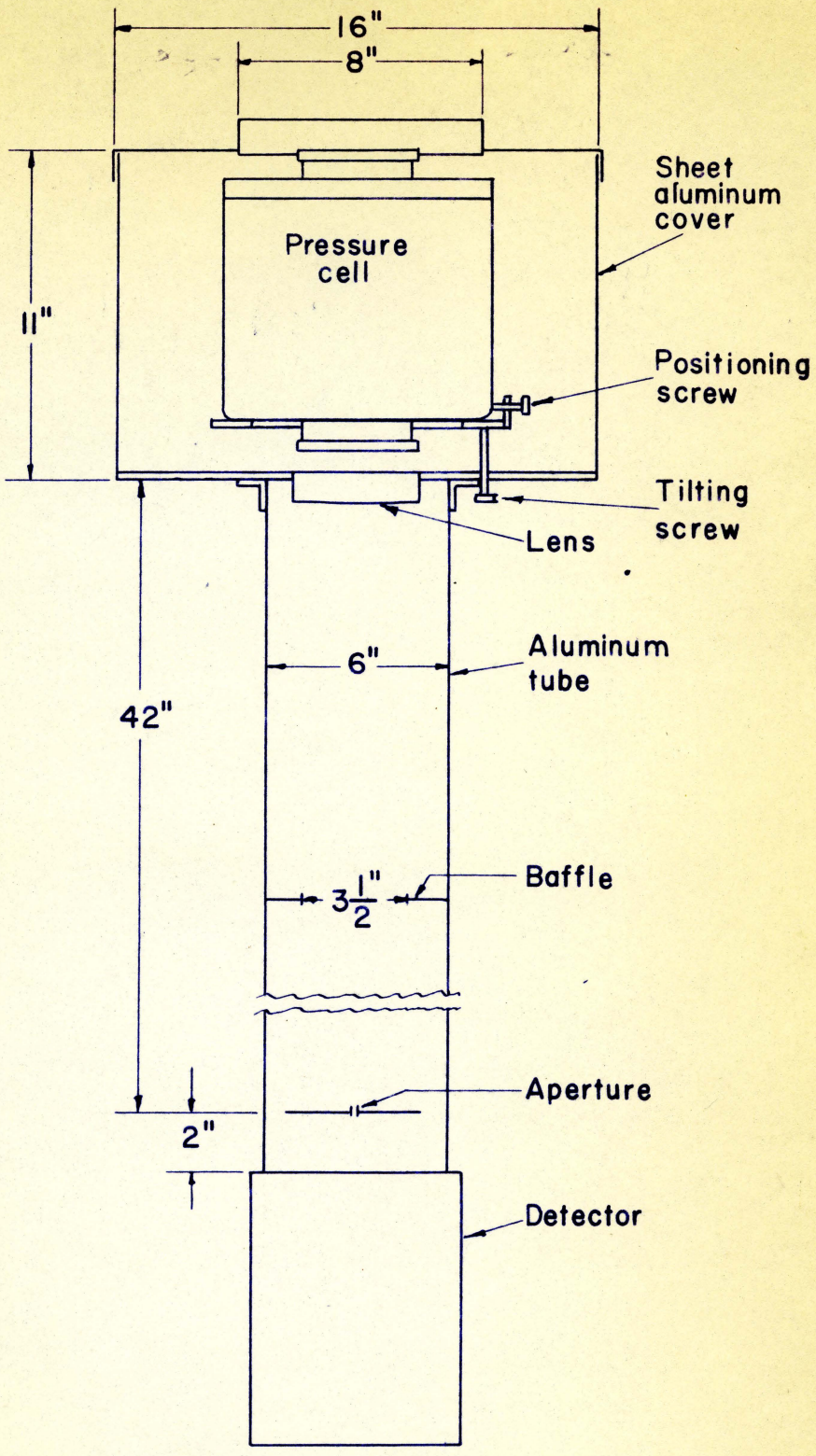
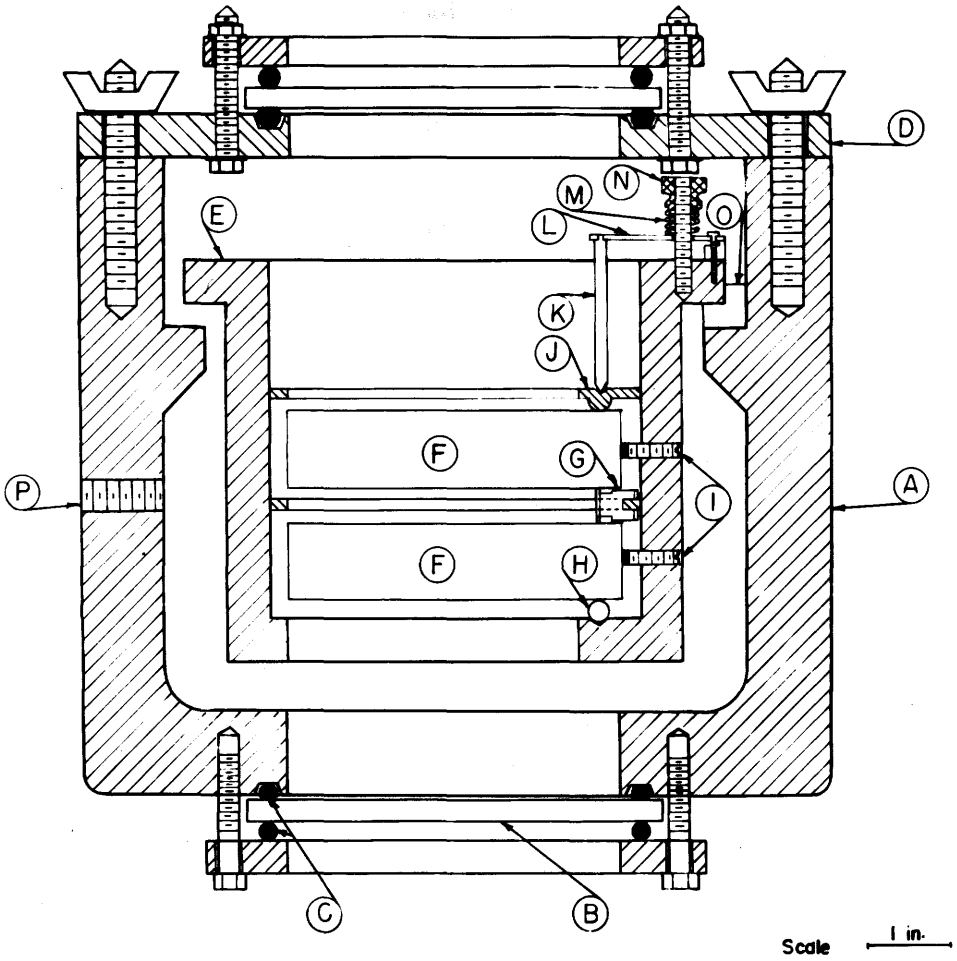

16"
8"
11"
Pressure cell
Sheet aluminum cover
Positioning screw
Tilting screw
Lens
6"
Aluminum tube
42"
3 1/2"
Baffle
Aperture
2"
Detector

FIGURE 3-1a



A - aluminum pressure vessel
B - windows
C - O rings
D - cover plate
E - steel plate holder
F - Fabry-Perot plates (4" Dia)
G - 1 cm. spacer (invar)(3)
H - 1/4" steel ball
I - position screws (6)
J - pressure ring
K - pressure rod (3)
L - leaf spring (3)
M - coil spring (3)
N - pressure nut (3)
O - poly styrene support (3)
P - pressure inlet

FIGURE 3.1b

Mounted on the same platform as the pressure cell are the valves and orifice that regulate the pressure in the cell. The valves are electrically operated so can be run automatically by the mercury manometer which makes contact at the highest and lowest pressures required controlling a relay which in turn controls the valves.

The objective lens is mounted in the base plate to make sure its axis is parallel and concentric with the apparatus. The tube which acts as a light tight connector between the lens and the aperture and detector is screwed solidly to the base plate. The base plate is spring mounted on the dexion stand and also serves as one end of a light shield that surrounds the pressure cell.

The whole apparatus, including electronics, is in the North Penthouse of the Physics Building. There was an 8 inch diameter hole cut in the roof with a cover so that a movable mirror can be placed on top to bring in the desired light. The distance between the instrument and the roof is 10 inches and it is about one inch off the floor.

3.2 The Etalon

The Fabry-Perot plates had been obtained before the author undertook the project as they must be ordered about a year in advance of when they are required. The plates are 4 inches in diameter, 0.91 inches thick, and when they were received had a semi-reflecting aluminum coating on their flat surfaces. Shortly

after, they were coated with a nine layer zinc sulfide and cryolite dielectric for testing purposes as will be discussed in Chapter IV.

The plates must be held in a fixed position with respect to each other, so that the adjacent faces are parallel and a given distance apart. There are many ways one could visualize spacing them but the simplest method seems to be satisfactory, that is, to put a spacer between them and hold them against it.

Now, the orientation of the plates must be considered. If there is going to be much auxiliary equipment such as prism or grating monochromators preceding the Fabry-Perot, then perhaps there may be reason for standing the plates on edge so the optical axis would be horizontal. Since the primary purpose of the instrument is to look at the sky, it is easiest to get the required light in if it is mounted vertically.

The plates are resting one above the other with the spacer in between and because they are horizontal there is no pressure required to hold them in place. There are positioning screws in the etalon which prevent the plates from sliding around when lifting them out of the instrument but they apply no forces on the plate as they are just touching.

The spacer is second in order of importance only to the quartz plates themselves. This is so because if the plates are not parallel it is useless to have good quality plates. The spacer is usually made of either quartz or invar because of their low

coefficient of expansion. Since quartz spacers are expensive to buy, and it is difficult to know in advance what spacing will be required, it seemed worthwhile to make a simple spacer using invar.

A suitable spacing distance t, had to be determined. The plates had not yet been tested, as a spacer is required for testing, so two values for $N_\mathcal{E}$, the finesse of the etalon, were chosen, 25 and 50. Since the auroral green line (5577 A) was going to be the first thing studied, the instrumental half-width would have to be the same order of magnitude or narrower than the half-width of the green line which is approximately .05 cm^{-1}. Two values were assumed for the half-width of the etalon function ϕ, namely, .05 cm^{-1} and .01 cm^{-1}.

From the theory we have $\Delta\sigma = \frac{1}{2t}$, $e = \frac{\Delta\sigma}{N_\mathcal{E}}$ which defines $t = \frac{1}{2eN_\mathcal{E}}$. The values of t and $\Delta\sigma$ using the various combinations of e and $N_\mathcal{E}$ given in Table 3.1 allow us to chose a spacing for the plates.

TABLE 3.1

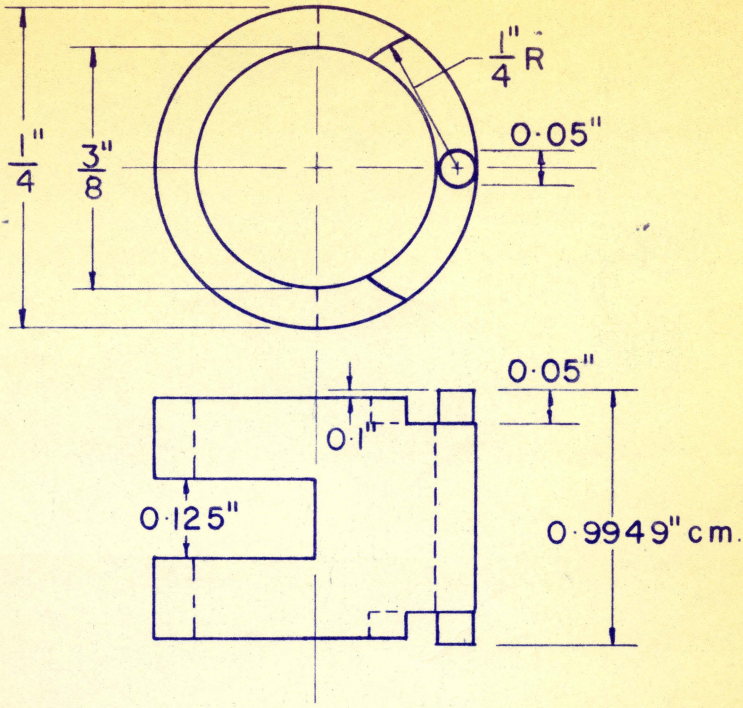
$N_\mathcal{E}$	$e(cm^{-1})$	t(cm)	$\Delta\sigma$ (cm^{-1})
25	.05	0.4	1.25
25	.01	2.0	0.25
50	.05	0.2	2.5
50	.01	1.0	0.5

The value t = 1 cm was chosen because it seemed to give enough distance between orders, $\Delta\sigma = 0.5$ cm^{-1}, and e would probably lie between .01 and .02 for $N_\mathcal{E}$ being 50 and 25 respectively which

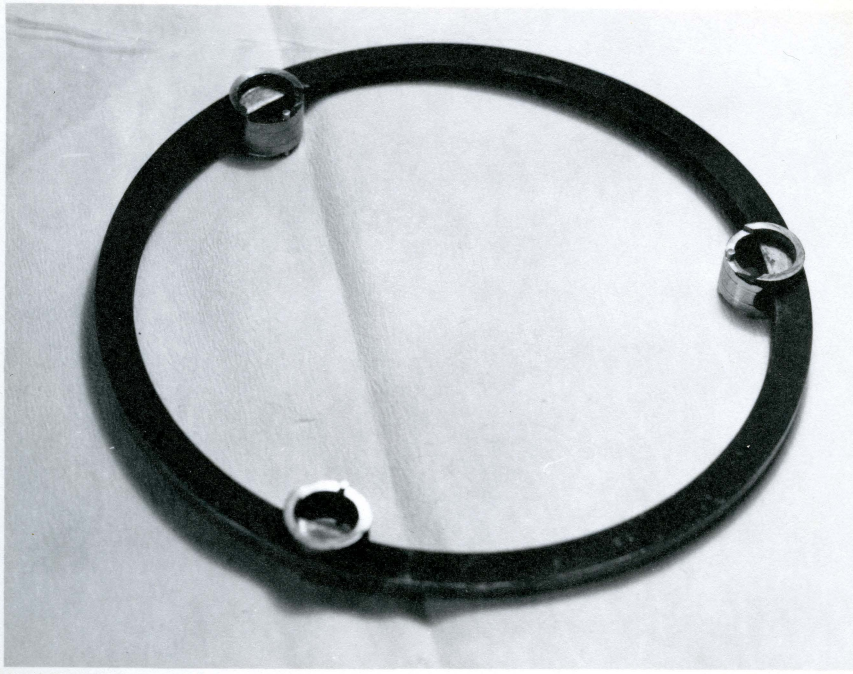
should give enough resolution.

It is not feasible to make a spacer so accurate that once ground to the proper length it would keep the plates in adjustment and give the same results if removed and then replaced. There is obviously need for some means of adjusting the plates once they are assembled with the spacer between them. This is done by having only three small points of contact where the plate touches the spacer. Pressure on the top plate above the point of contact compresses the spacer slightly allowing adjustments to the parallelism to be made. The pairs of contact points where the projections on the spacer touch the upper and lower plates, are spaced 120° apart and each pair is parallel to the optical axis of the plates. In other words, one is directly above the other.

The spacer that was designed and made is shown in Figure 3.2. Meissner (1941) states that an area of contact of 1 to 1.5 mm square is satisfactory for the dimensions of the projections. The only invar on hand was $\frac{1}{2}$ inch diameter tubing with 1/16 inch wall thickness. The fact that it was thick walled tubing made it possible to mount it off center in a lathe so that a projection .05 inches in diameter could be turned out. This gives an area of 1.27 square mm which would agree with that given above. The length of the projection was calculated using Young's modulus. By assuming $\frac{1}{4}$ pound ~~is~~ a reasonable force to apply to compress the spacer one half wavelength of sodium light (one order), the length



INVAR SPACER



SPACERS MOUNTED IN RING

FIGURE 3·2

of the projection becomes 0.05 inches. The plates weigh over a pound each themselves, which means they should be in good contact, Hence the adjustment will be smooth. Each spacer had a slot milled in it perpendicular to the center line of the projections so that they could be soldered in a brass ring with a separation of 120°. Brass was used because it was available, and the spacer was designed so that if there was any change in dimension in the brass due to temperature it could be hardly transmitted to the actual spacing distance between the projections. The whole etalon being inside a temperature controlled cell further ensures the stability of the spacing distance.

There are two schools of thought on how spacers should be ground. One is that the spacer should be optically flat at the surface of contact as described by Meissner (1941) and the other is that the projections should be rounded as for example Wright and Curtis (1931). It was first attempted to grind the spacers flat. This was done by grinding one projection on plate glass while the other two rode on oil also on the plate. They were rough ground until they were the same length as far as could be ascertained with a micrometer. A more accurate measurement of the relative lengths of the three spacers is done by putting the plates and spacer in the holder and counting the change in order of the Fabry-Perot rings as the eye is moved from one spacer to the other. If for example when moving from spacer number 1 to

number 2, the Fabry-Perot rings grow in diameter and four rings appear, and when moving from 1 to 3 two new rings appear, then number 2 is four orders (1.2 microns) high, and number 3 is two orders (0.6 microns) high, with respect to number 1. This is the way the spacer appeared when tried the first time after rough grinding which means it is easy to get fringes to appear when the spacer is made while checking carefully with a micrometer. When, however, the spacers were to be ground more closely it was very difficult to get reproducible results. When using quartz spacers as described by Meissner (1941) they are ground optically flat and tested by looking at the localized fringes where the spacer contacts the plate. Invar does not polish as nicely as quartz and fringes could not be seen to make sure they were in good contact with the plates. Thus in an ordinary room where dust invariably settles on the plates and spacer, and hence changes the spacing, it is difficult to get reproducible results. If rounded projections were used it would seem the chance of getting good contact with the plates would be much greater.

When it was realized that grinding by hand with fine grinding compound to get a nice finish was too slow, a small polisher was made out of an old timing motor. Crocus cloth mounted on a cork disc worked quickly yet left a smooth finish. The spacer was moved around when polishing so as to get a rounded surface. It was possible to remove about one order from the spacer when polished for one second on each end. The end result was that the

three spacers were the same length to less than one order (about 0.3 microns). This last bit is easily adjustable by light spring pressure on the plates as discussed in the next section on the holder and complete etalon.

The spacers have been very satisfactory since the etalon will stay in adjustment for a week (it has never been given a chance to go longer) when in the temperature controlled cell.

The etalon, Figure 3.3, was designed with spacers up to 4 cm in mind. All that has to be done to accommodate a new spacer is to make a different set of pressure rods and perhaps new positioning screws. Another thing that helped to determine the size of the holder was the fact that since the refractive index method of scanning is being used it is necessary to allow space for the gas to flow between the plates. A minimum distance of $\frac{1}{4}$ inch between parts through which the gas must flow was imposed to keep turbulence and pressure gradients to a minimum.

The bottom plate is in solid contact with the three ball bearings which sit in conical recesses in the bottom of the holder. The positioning screws do not support it at all. What must be done then to adjust the parallelism of the plates is to compress one or two of the spacers slightly. The force is applied bymeans of the adjusting nut which compresses the coil spring. The coil spring can be replaced by one stiffer or weaker depending upon how much force is required for adjusting. The leaf spring acts as a

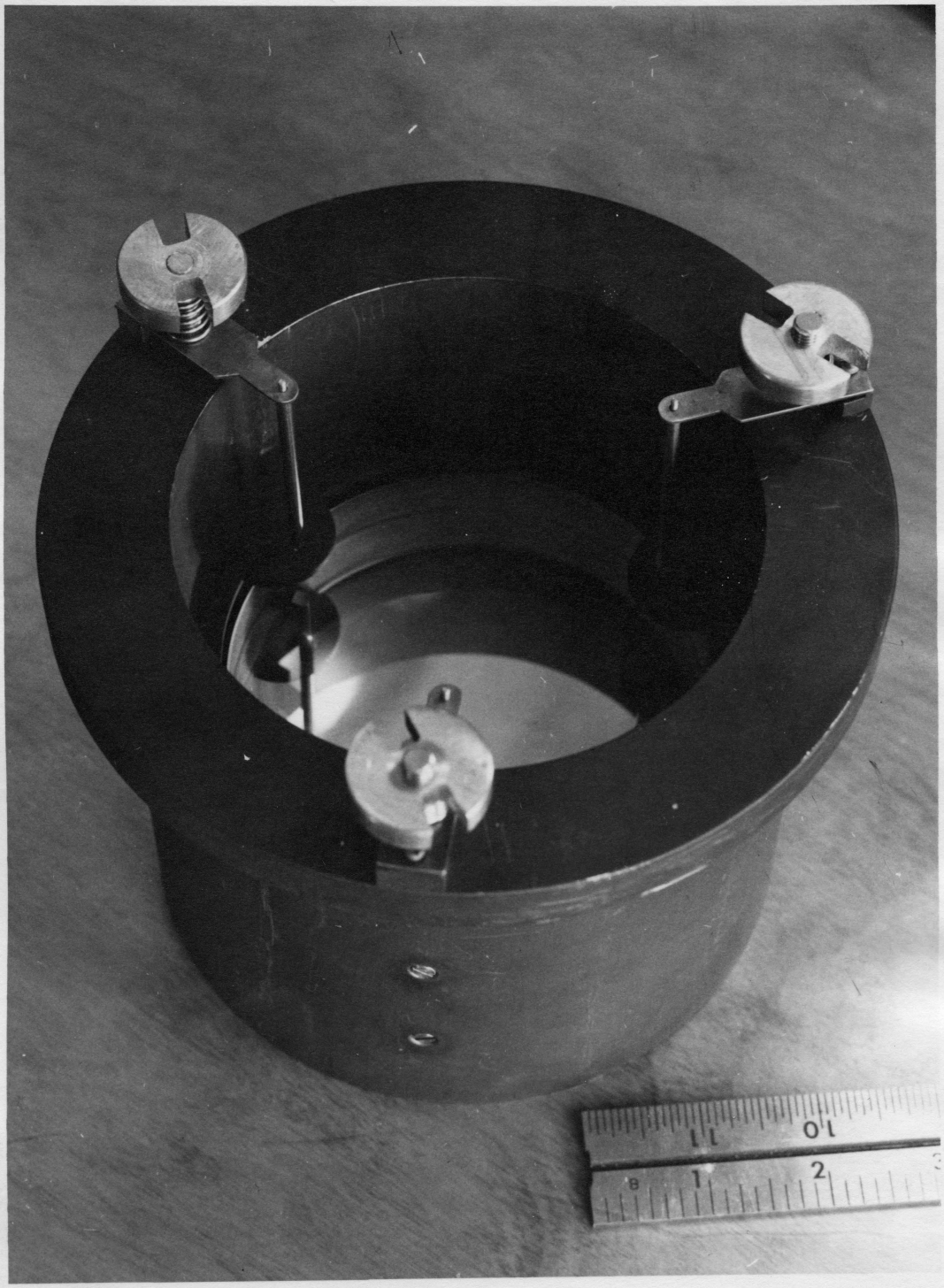


FIGURE 3.3

See Also FIGURE 3.1b

lever to transmit the force to the pressure rod. It is very important to have the ball bearings, the spacers and the pressure rods all in line so no torques on the plates are developed which would twist the plate as well as compress the spacer. Hence the pressure ring aligns the pressure rods so that the force is in the required place. Since the bottom plate is very steady, the force is transmitted from the pressure ring through the top plate to the spacer which is compressed the desired amount, bringing the plates in adjustment.

The holder was machined from a bar of 1.3% nickel steel (SPS - 245 similar to SAE - 3140). This steel is in a state which should hold creep to a minimum. It was rough turned to 1/16 inch oversize in all dimensions and then stress relieved in an electric oven. (1) After cooling it was turned down to the required size taking small cuts so that it would still be in a stress relieved state.

An accessory that had to be made to ensure safe handling of the plates was the plate carrier shown in Figure 3.4. This has three stainless steel legs that can be rotated and locked in two positions by means of the long nuts which are threaded to the legs.

(1) The turning was done by J. Roney in the shop of the Department of Mechanical Engineering, U of S. It was stress relieved for 4 hours at 350 °C, and cooled overnight in an oven supplied by Dr. R. L. Eager of the Chemistry Department, U of S.

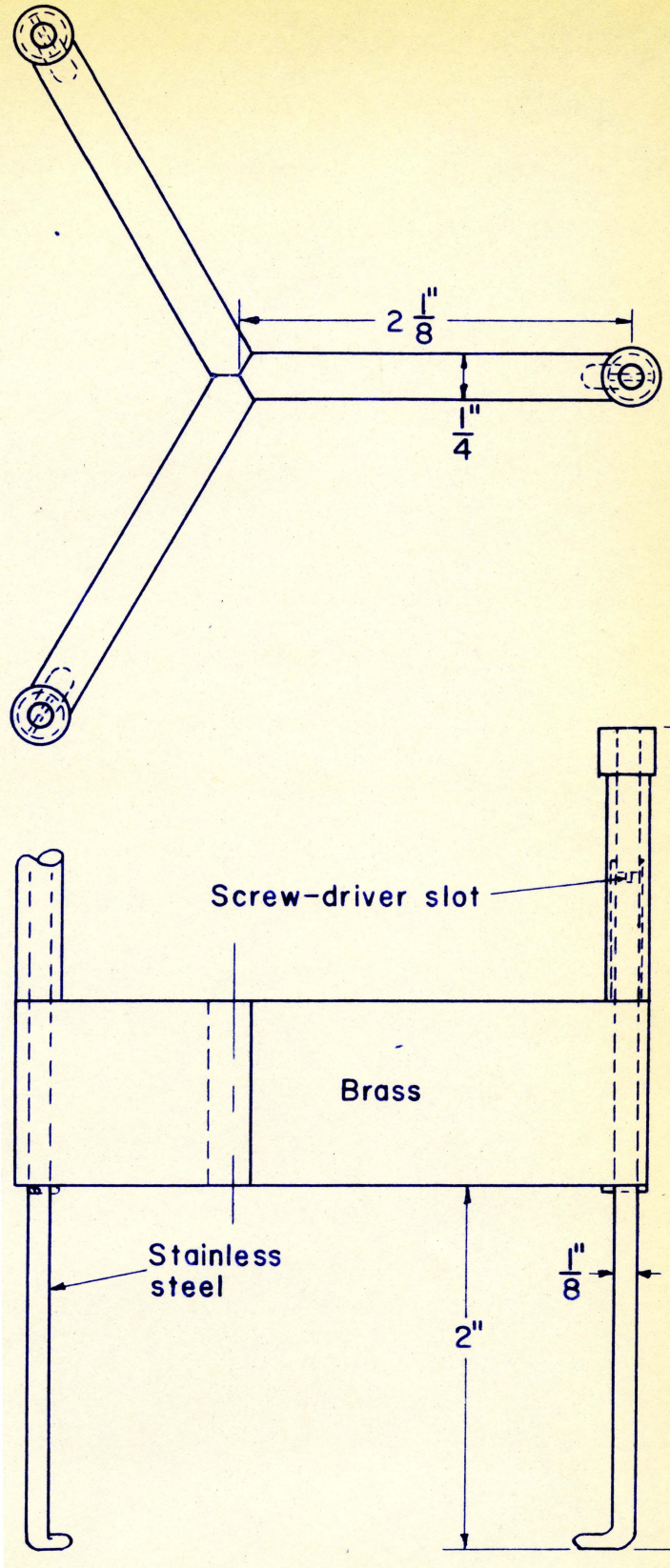


FIGURE 3·4

This permits the feet to be positioned under a plate making it possible to lift it. When the plates have to be cleaned, in order to coat them with dielectric multilayers, they have to be handled a great deal and immersed in acids (nitric for removing the dielectric) and bases (sodium hydroxide for removing aluminum) as well as soapy water and alcohol. Stainless steel seemed to be the proper material for the legs, and rather than try to glue some soft material that could stand the rough treatment onto the feet, they were carefully polished smooth, and round so they could not scratch the plates. The reason the feet have to be thin is because the distance between the brass ring of the 1-cm spacer and the plate is about 1/8 of an inch making it even more difficult to find a suitable soft covering.

The assembly of the etalon is done in the following manner. The ball bearings are put in their conical holes. The bottom plate is carefully lowered in with the plate carrier which is then removed and the plate positioned with the three positioning screws. The spacer is lowered in with wire hooks and positioned. The top plate is then lowered in with the carrier. It is advisable to check the adjustment at this point because if there is any dust under the spacer, the top plate and then the spacer if necessary can be easily removed. To be adjusted easily the spacers should be within one order of each other. If they are not quite, this can be remedied by moving the appropriate spacers onto the

dielectric from the clear spot. In this manner very close adjustment can be realized before applying any pressure. The reason that one spacer can be moved onto the dielectric and not the others is that the clear spots caused by the holders that hold the plates in the evaporating chamber are not exactly 120° apart. One must always remember to make sure that the spacer is directly above the ball bearing. There is now another reason for having a round spacer rather than optically flat at the contact point because flat ones would not be contacting the plate as they should if they were part way up the dielectric. The pressure ring is then lowered in with the rounded spots over the spacers. The leaf springs are put on and then a heavy handle which fits over the adjusting studs is put on using the adjusting screws. This allows the etalon to be carefully lowered into the pressure cell. The handle is removed and the coil springs, adjusting nuts and finally the pressure rods are put in place.

3.3 Pressure Cell and Temperature Controller

The purpose of the pressure cell is to support the etalon, to allow a change in pressure of the gas surrounding the plates to be made, and to provide a temperature controlled environment for the etalon. Because of its large size as shown by Figure 3.5 and the need for good heat conduction, cast aluminum was used.

A minimum clearance of $\frac{1}{4}$ inch between the etalon and the cell was adhered to in order to make room for the free passage of gas

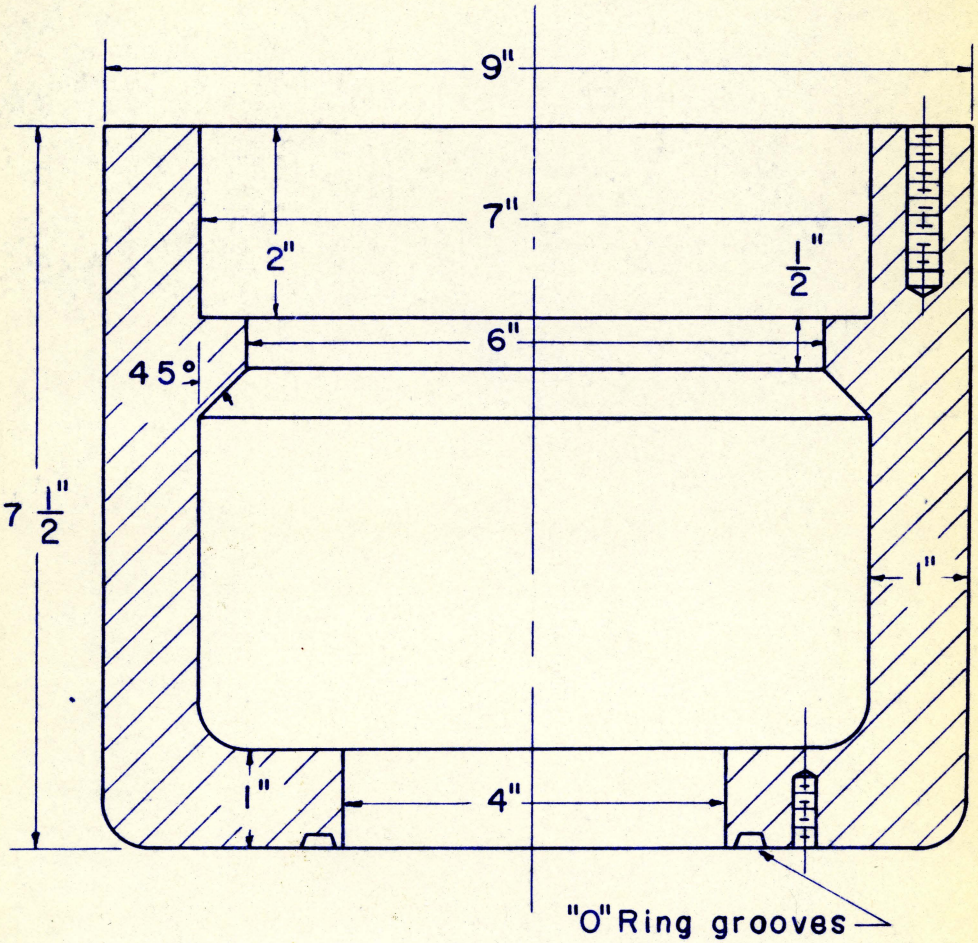


FIGURE 3·5
See Also
FIGURES 3·1 & 3·12

in the cell. The etalon rests upon three lucite blocks that keep it on center but do not permit much heat conduction. There is thus no metallic contact from the cell to the etalon. This is desired so that the temperature variations of the cell will be reduced in magnitude in the etalon. The lid rests upon a rubber gasket and is held down with 12 wing-nuts on 3/8 inch studs that are threaded into the casting. Coarse thread studs were used throughout when threaded into the aluminum because the aluminum does not wear well.

The lid, made of $\frac{1}{2}$ inch steel plate, has a 4 inch diameter hole in the center for the light to pass through and is machined for an "O" ring so that a window may be sealed to it. The bottom of the casting has a similar "O" ring groove cut in it for the lower window. The "O" ring grooves are cut so that when the "O" ring is compressed properly the glass window will be 1/16 inch from the metal, which makes it possible to get the proper compression without clamping the glass to the metal. There is another "O" ring but no "O" ring groove under the retainer ring to cushion the glass. The force is applied by the 3/8 inch steel retainer ring which is the same both top and bottom. There are 12 $\frac{1}{4}$ inch coarse thread studs in the aluminum in the bottom to hold the retainer ring, and $\frac{1}{4}$ inch fine thread bolts threaded into the steel cover plate, with copper washers underneath the heads for a vacuum seal, to hold the retainer ring at the top.

The windows were made from standard $\frac{1}{4}$ inch plate glass and are 5 inches in diameter. From information obtained from Libby Owens Ford Glass Company the modulus of rupture of plate glass is 6,500 psi. Corning gives the modulus of rupture for pyrex as 10,000 psi.

The formula $P = \frac{3.3\ M\ t^2}{A\ S}$ was used (from Libby Owens Ford) where P = pressure is psi, M = modulus of rupture, t = thickness, A = area and S = safety factor. Assuming P = 14.7 psi, M = 6,500, t = $\frac{1}{4}$ inch and A = $\frac{\pi\ 4 \cdot 5^2}{4}$ square inches, this gives S = 5.7. This safety factor is hardly large enough. Libby Owens Ford suggest a value between 5 and 10. If 3/8 inch glass were used then S = 12.6 which is better, so if the cell were to be used at pressures up to one atmosphere as a regular thing, 3/8 inches thick would be best. However, using high resolution it is only necessary to go over at most 10 cm Hg which is a reasonable pressure range.

The cell was bored and tapped for 3/8 inch pipe thread so that a 3/8 inch pipe to $\frac{1}{2}$ inch refrigeration fitting could be used to bring in the gas. The complete pressure cell was tested under water and was found to leak like a seive. The air was going through the one inch thick aluminum casting. The inside of the casting was painted using glyptol and then put under pressure to force the paint into the holes and also see where the holes were. This was dried using a heat lamp and repeated enough times so that nearly all the holes were filled. The apparatus

holds the pressure well enough now. The outside was also painted with clear airplane dope as a safety measure and as insulation for the windings. The complete assembly was tested at 30 psi.

A mechanism that would allow the instrument to be adjusted from outside the pressure cell was considered but rejected as unnecessary with the one centimeter spacers as they stay in good adjustment. Later however, when testing at low order it was found impossible to remove the lid, adjust it, and then put the lid back without affecting the adjustment. It was necessary to design a mechanism that could be used to turn the adjusting nuts from the outside.

Figure 3.6 shows the modifications carried out. The adjusting nuts were replaced by larger ones that had slots in which the fork of the adjusting rod fit. These slots were made extra large so that once the adjustment is made the forks can be turned such that they do not touch the etalon. The $\frac{1}{4}$ inch diameter steel adjusting rods have an "O" ring groove cut in them that seals in the $\frac{1}{4}$ inch diameter reamed hole in the $\frac{1}{2}$ inch thick lid. The nuts on top are removable so that when the pin above the lid is removed the whole rod can be pulled out. These outside adjusters are very useful and are valuable even when using the one centimeter spacers, as the adjustment can be made more exactly when the lid is bolted down and no more jarring will take place. The sheet aluminum box surrounding the cell has a hole in the top

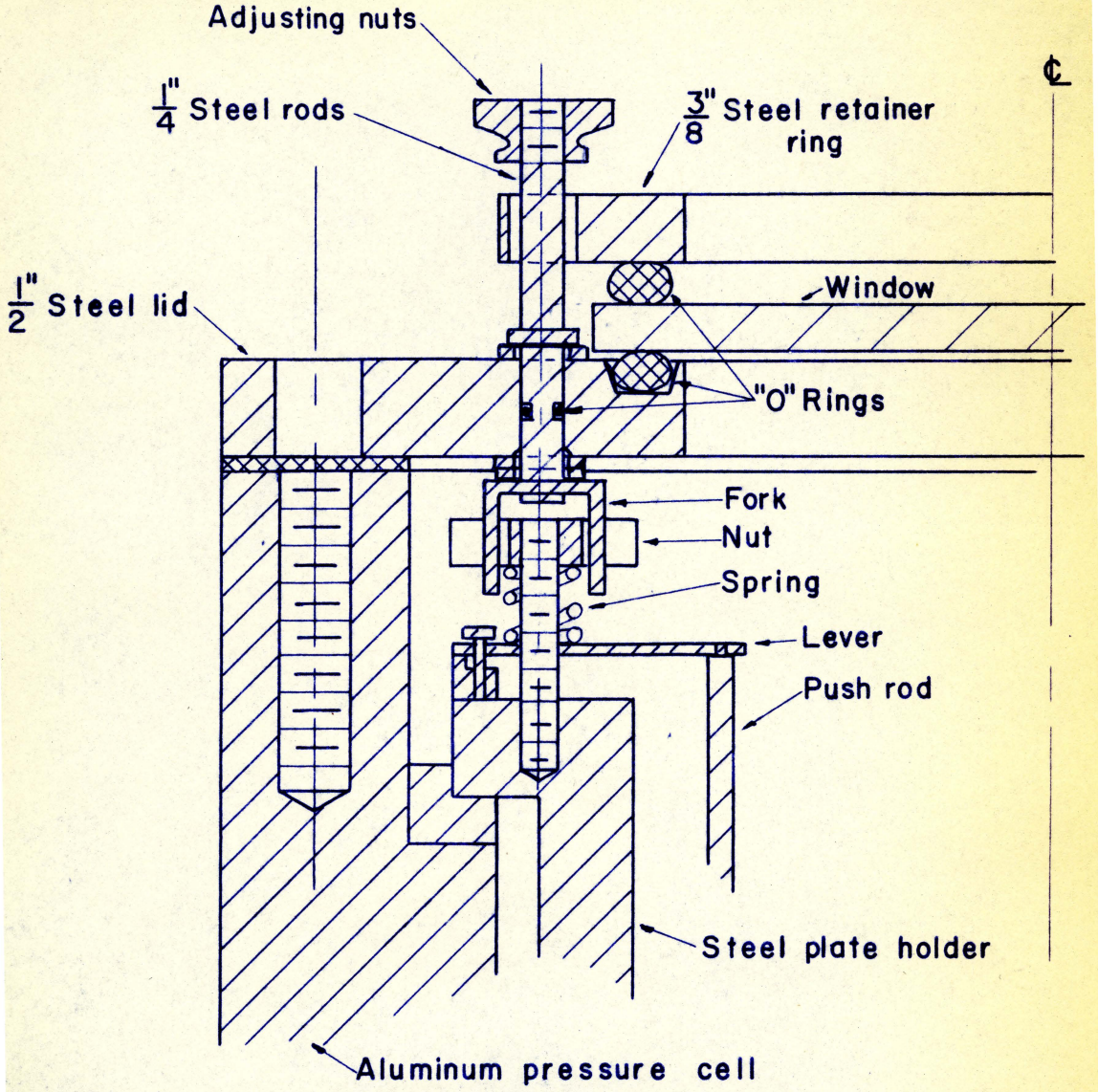

Adjusting nuts
1/4" Steel rods
3/8" Steel retainer ring
1/2" Steel lid
Window
"O" Rings
Fork
Nut
Spring
Lever
Push rod
Steel plate holder
Aluminum pressure cell

FIGURE 3·6

large enough so that the adjusters can be reached when the whole instrument is assembled. There are no leaks that could be detected from the "O" ring seals in the adjusters so they appear very satisfactory.

The apparatus had to be temperature controlled, especially when the original idea was to have it outdoors. The pressure cell was wound with four separate coils of wire so that they could be hooked up in different combinations from parallel which gives 413 watts, to series which gives 26 watts. Nichrome wire having resistance of 1.3 ohms/foot was used and when operated on 110 V A C was drawing nearly one amp of current per coil. For detecting the temperature the pressure cell was wrapped with Balco Resistance Wire which is 23.6 ohms/foot. 1,000 ohms were required making 18 turns which were split up into three separate coils. The temperature sensing coils were wrapped close to the heating coils to minimize the overshoot caused by slow heat travel.

The temperature control seems to be good as the heating current is turned on and off every four minutes or less, when operating at 26 watts in a heated room with the instrument 15 degrees Fahrenheit above room temperature. The important thing however, is that the plates stay in adjustment.

The electronics are built around a Honeywell Moduflow control unit. This was used as shown in Figure 3.7. The terminals indicated as being in Moduflow are in the control unit which is

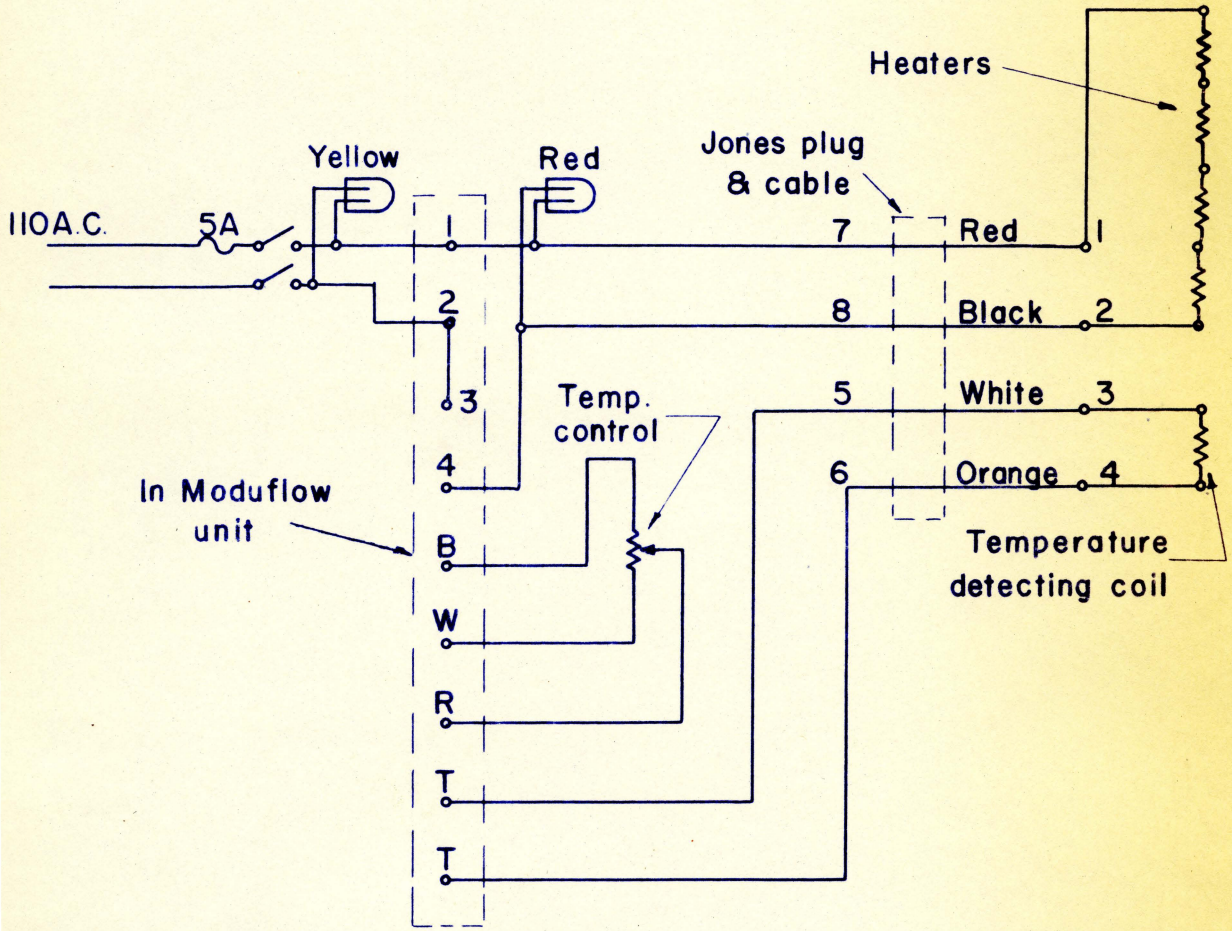

Heaters
Yellow
Red
Jones plug
& cable
110A.C.
5A
1
2
3
4
B
W
R
T
T
7
Red
1
8
Black
2
5
White
3
6
Orange
4
Temp.
control
In Moduflow
unit
Temperature
detecting coil

FIGURE 3·7

mounted on the back of the control panel. There is a main power switch together with pilot light, a red light (to indicate when the heating current is on), and temperature control rheostat mounted on the front of the control panel. An 8 terminal Jones plug and 8 wire cable is shared with the pressure regulator, described in the next section, which goes to the main terminal block mounted on the back of the base plate of the instrument. Another terminal block is provided for connecting the heaters in the desired configuration to get the proper wattage.

3.4 Pressure System and Scanning

In a spectrometer one would like a linear change of wavelength with time so that the chart recorder, which runs linear with time, could be easily calibrated. It was shown in the theory that refractive index scanning would be desirable when working at high order, so we shall be concerned only with this method.

Differentiating $m_o = 2 n \sigma t$ keeping the order m_o and the plate separation t constant, gives $\frac{d\sigma}{\sigma} = -\frac{dn}{n}$. Since both $\sigma >> d\sigma$ ($\sigma = 17{,}900$ and $d\sigma \doteq 1$) and $n >> d n$ ($n = 1.0011$ and $d n \doteq .0007$) when scanning over a few orders, to a very good approximation, $d\sigma = a\, d n$ where a = constant.

For constant temperature $n - 1 = b p$ where b = constant and p = pressure, thus $d n = + b\, d p$. Combining with the above gives $d\sigma =$ constant X d p. Thus what is required to give a linear change in σ with time is a linear change in pressure with time.

There are many different methods for obtaining a linear variation in pressure. Jmaeff (1959) describes a few methods based on servomechanisms. These are sound in principle but are quite complicated. A simpler method suggested by Rank (1954) was tried. This depends upon having high speed flow of gas through a small orifice. Rank found that the mass of gas flowing through a small orifice, if at sonic velocities, is nearly constant and independent of back pressure in the cylinder to which it flows. If there is constant mass flow to a vessel of volume V, then $p = \frac{nRT}{MV}$ and $dp = \frac{dm\,RT}{MV}$ if T is a constant. Thus if $\frac{dm}{dt}$ = constant, then $\frac{dp}{dt}$ = constant.

The orifices were made by letting pyrex capillary tubing contract while heating a small length to the softening point. This method is very satisfactory and fine orifices can be made easily, except a few had to be made before getting one the desired size. The cross section of a good tube would be similar to Figure 3.8.

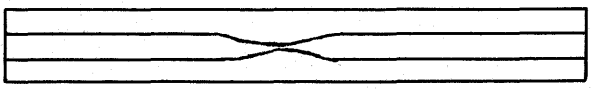


FIGURE 3.8

A typical example is a tube with inside diameter of orifice 0.22 mm (# 7 tube). With this tube in the apparatus σ changes by one wavenumber in 12 seconds. Tube # 5 with inside diameter about 0.1 mm changes σ one wavenumber in about 60 seconds. The ratio of the area is about 5 and the ratio of their speeds also about 5. The reason for the approximate nature of the diameter is because they were measured by looking down the tube with a long

focus microscope and the exact minimum diameter was not assured since some of the fine orifices are slightly curved. The length of orifice may also be important as some tubes that have a gradual slope to a small constriction do not seem to be as linear as the ones with quick change to a small constriction.

The tubes were tested using air for the high pressure and discharged into a vessel equipped with a manometer. They were thought to be satisfactory but much better tests were possible once the apparatus was assembled. A graph of order against time is shown in Figure 3.9. This graph was made using the results obtained when running over a large pressure range at high order which goes through a large number of orders.

Since $m_o = 2\ n\ t\ \sigma$ the change in the order m_o should be linear with n and hence with time. The change in order over the scan is found easily because the line used was monochromatic (Kr 5570 A) and the peaks denote the orders. The top curve which is the actual measured values of order against time is very straight for the first 24 orders which corresponds to an absolute pressure change in the vessel from 0 to 58 cm Hg. The other curve is plotted using the formula $m = b\ (\ 1 - e^{-at}\ )$ which is the curve obtained when a vessel of constant pressure is flowing into a vessel of constant volume through a theoretical sub-sonic orifice that has the same initial slope for order plotted against time and the same pressure on the high side. The actual curve is

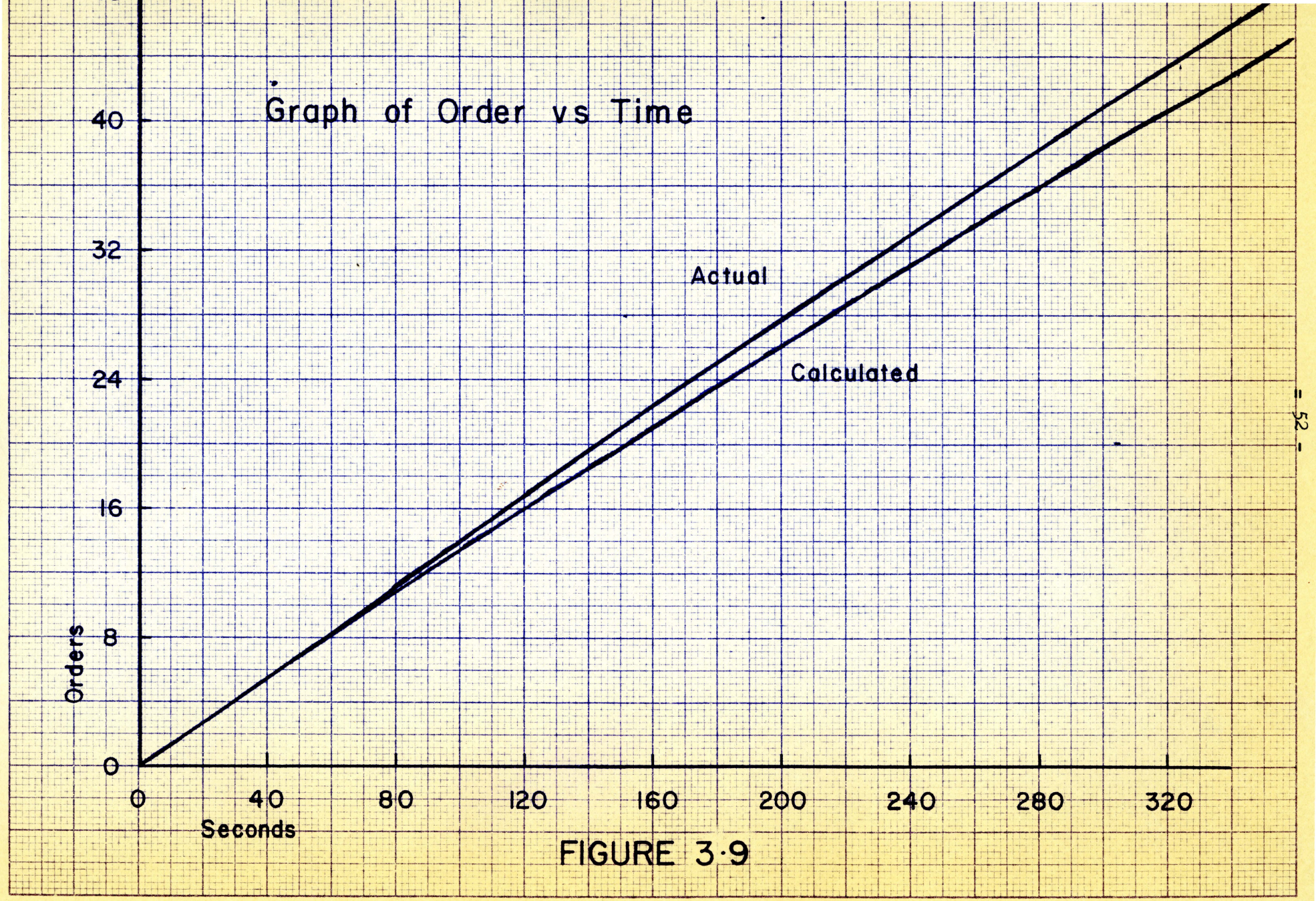


FIGURE 3·9

much more linear than the theoretical and indicates that sonic or near sonic flow is achieved. When operating over a small region of a few orders as is usually done it is easily seen that it will be very linear.

It was found in the preliminary tests with air that a pressure of 60 psi or more would be required to get linear flow. A curve of rate of change of pressure in the vessel at atmospheric was plotted against high pressure to the orifice and found to be linear from 10 psi to 60 psi, the rate of change of pressure increasing four times over this pressure range. This means the high pressure must be constant to get constant mass flow.

University compressed air line pressure was found to vary from 60 to 80 psi and was also quite dirty so was not considered beyond the testing stage. Another source of high pressure gas is bottled gases. If they are liquified and have the proper vapour pressure at room temperature any kind of regulating can be dispensed with since the only changes in pressure are due to temperature, which are relatively slow. An important advantage of using a gas other than air is that its refractive index n, may be larger. This means that since $d\sigma$ = Constant X d n and d n = Constant d p, the pressure change for a given d σ can be less using a higher refractive index. The International Critical Tables give some indices of refraction and others can be calculated approximately using a form of the Clausius-Mosotti formula as given by

Glasstone (1940). $[R] = \frac{(n^2-1)M}{(n^2+2)\rho}$ where R is determined for each gas by consideration of the components and bonds in the molecule. M is the molecular weight and ρ the density. For a perfect gas $\frac{M}{\rho}$ is a constant, 2.24×10^4 cm /gm and the formula can be written $n^2 = 1 + \frac{3R}{\frac{M}{\rho} - R}$. Let us take for example Freon-12 (difluorodichloromethane). The R value for C:Cl is 6.57 and C:F is 1.60, thus the total R is 2 X 6.57 + 2 X 1.60 which gives 16.34. This value of R gives n = 1.0011 which means if one used Freon-12 instead of air (n = 1.00028) it should be possible to scan through about four times the wavelength interval using the same pressure range.

Freon-12 was chosen because it has a vapor pressure at room temperature of about 70 psi, and is available commercially in a pure form since it is used in refrigeration. If one tries to get a higher index of refraction the vapor pressure is usually less so Freon-12 seems to be a happy choice. Propane or Butane would also be useful, having the pressure and refractive index similar to Freon-12, but they are not locally available in a pure form.

The way in which the gas pressure is utilized is to let in some gas through the orifice until the desired pressure is reached, bring down to atmospheric (or vacuum if a larger range is required) and then repeat. The graph of pressure against time should be a sawtooth function with a slow linear rise and a fast drop. The final form of the gas handling system is shown in Figure 3.10.

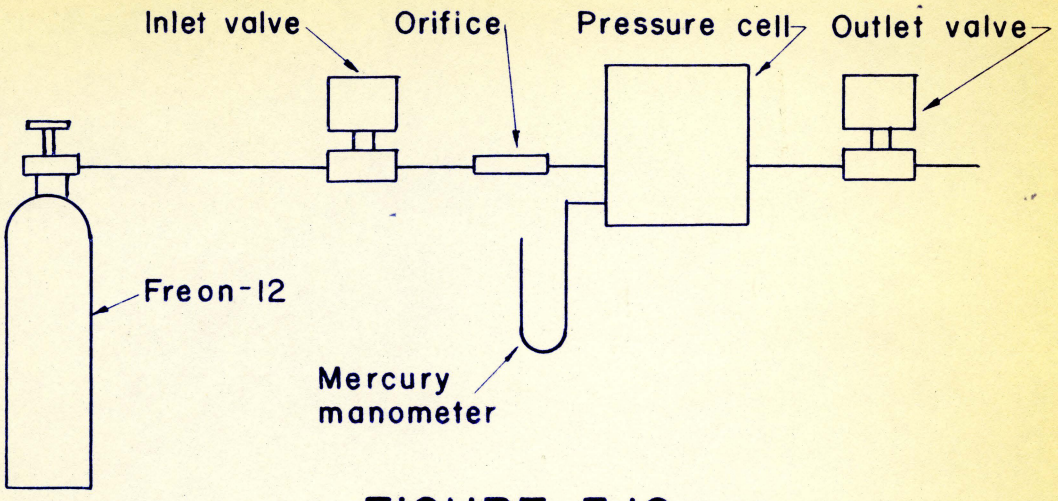


FIGURE 3·10

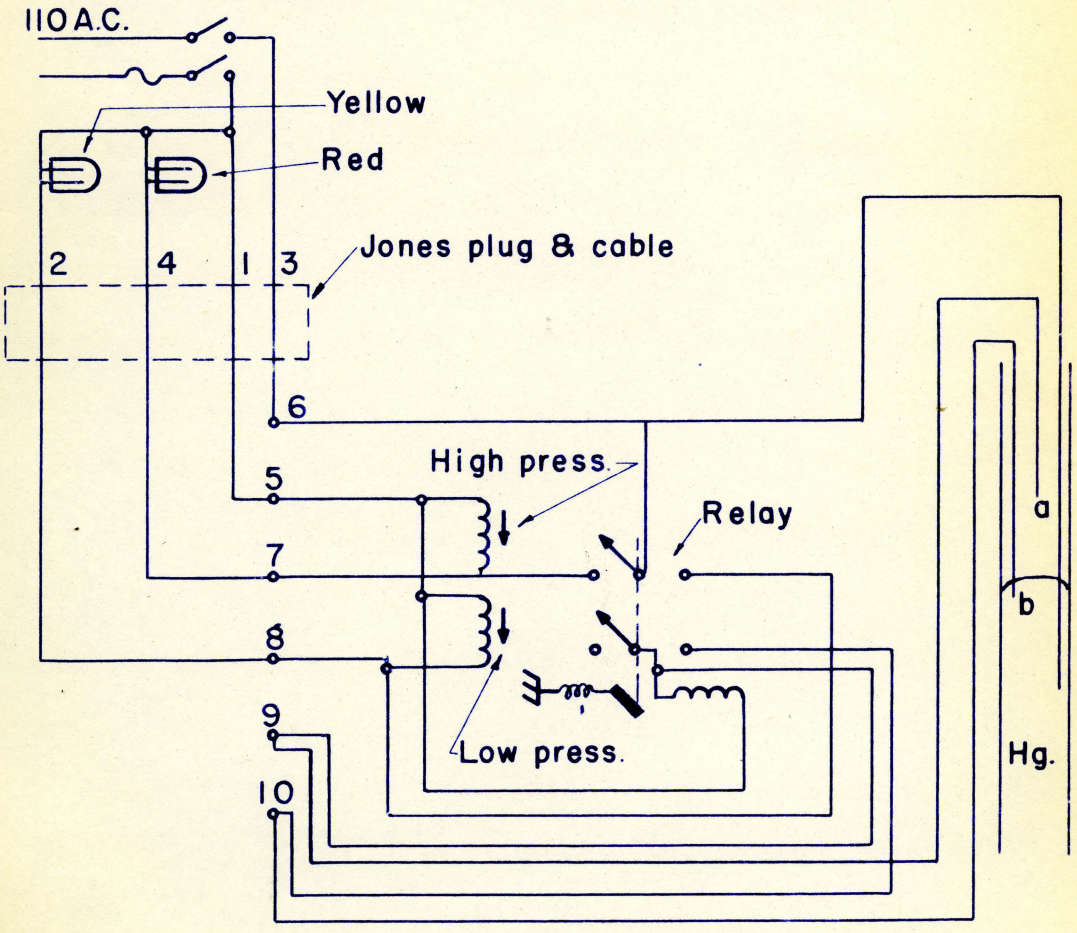


FIGURE 3·11

The high pressure is brought to the inlet valve with $\frac{1}{4}$ inch copper tubing. When a scan is required the valve is opened electrically and the gas flows through the orifice raising the pressure and the mercury in the manometer until the desired pressure is reached. There is a contact (a) in the mercury column that operates a relay to open the outlet valve and close the inlet valve when the mercury touches it (see Figure 3.11). The valves stay in this position until the mercury falls below another contact (b) which then operates the relay to open the inlet and close the outlet valve. This keeps repeating until the main switch is opened. Whenever the switch is closed the system is in the pressure rising state since the other state is realized only when the mercury touches the top contact. The system works well, except, the mercury gets dirty from sparking. If a lower voltage and a more sensitive relay were used or one stage of electronics, it would probably be more satisfactory. Another innovation that might be useful would be a system such that it takes a pulse from a push button to open the high pressure valve, then the system would go through one cycle and wait to be started again. This would be useful for long scans.

All the fittings and tubing for the apparatus are standard refrigeration $\frac{1}{4}$ and $\frac{1}{2}$ inch flare fittings modified to take "O" rings as described by Hodder (1958). These fittings when modified only require hand tightening and are very satisfactory for this

application where the pressure does not exceed 100 psi or get less than a few mm Hg. When it is required to scan over a larger wavelength interval than is possible using pressure only, a vacuum pump can be connected to the discharge line and a larger manometer used. In this way it is possible to scan over about 140 cm Hg corresponding to 88 orders at one centimeter spacing which is 34.4 cm^{-1} or 10.6 A at 5570 A. Of course this means that it can scan over 10.6 A at any plate separation as is explained in Chapter II.

3.5 Objective Lens and Aperture

The objective lens is required to focus the Fabry-Perot fringes, which are at infinity, onto the focal plane where the aperture is located.

Let the minimum diameter of aperture used be 1/8 inch (about 3.2 mm) since if it were much smaller there might be trouble in aligning the instrument and edge effects could become important. If one centimeter is assumed for the spacing between the plates then $\Delta\sigma = 0.5\ cm^{-1} = 500$ mk. A finesse of the aperture can be defined as $N_f = \frac{\Delta\sigma}{f}$ where f is the width of the rectangular aperture function. This is not a true finesse, since f is a constant depending on the angle θ subtended by the aperture at the objective lens, by $f = \frac{\sigma\theta^2}{2}$ as is shown in Chapter II, but it is a useful concept when the plate spacing is not being changed. Since the line to be studied will be around 50 mk wide a reasonable

half-width t, of the aperture function would be 25 mk which gives $N_F = 20$. Having said that the diameter d of the aperture is 1/8 inch and $f = 25$ mk, we have defined the focal length F of the lens. Using $f = \frac{\sigma\theta}{2}$ and $\theta = \frac{d}{2F}$, we obtain $F = \frac{d}{2}\sqrt{\frac{\sigma}{2f}} = 96$ cm. The lens purchased was a coated achromatic telescope objective with diameter 4 5/16 inches and measured focal length 105.8 cm.

The aperture (or exploratory diaphragm as it is sometimes called) is at the focal plane of the lens and is adjusted in size by removing and replacing it with the required one. Three apertures were made, with $\frac{1}{4}$ inch, 3/16 inch and 1/8 inch diameter reamed holes so that t can be found accurately. The values of t, θ and N_F are given in the Table 3.2 below for the lens described and spacing t = 0.9949 cm.

TABLE 3.2

Aperture (in)	θ (radians)	f (mk)	N_F
1/4	3.0×10^{-3}	80.8	6.22
3/16	2.25×10^{-3}	50.5	9.95
1/8	1.50×10^{-3}	20.2	24.85

The lens need not have a large clear field as would be necessary if photographing the rings, since only near axial rays are used, the angle θ for the $\frac{1}{4}$ inch aperture being only 0.17 degrees.

The platform which holds the aperture, ~~Figure 3.12,~~ can be

moved horizontally in two directions at right angles to each other by means of screws. This is necessary to make sure the center of the fringe pattern falls on the center of the aperture. The coarse adjustment as mentioned before is made by tilting the etalon which is done by tilting the whole pressure chamber. This leaves only very small adjustments to be made at the aperture. The whole platform is removable and can be replaced in the same position because it is on a 3 point kinematic mount.

3.6 The Detector and Electronics

The detector lens should gather all the light that goes through the etalon and aperture and focus it onto the cathode of the photomultiplier tube. If the image of the plates, formed by the objective lens, is focused onto the cathode of the photomultiplier by the detector lens, then all of the light that goes through the plates in the solid angle subtended by the detector lens from the plates will go to the cathode. Then all that is required is that the lens be large enough to include all the light that goes through the aperture and have a focal length such that the image of the plates on the cathode is of the required size.

A detector including lens, E.M.I. 9502 A photomultiplier tube in an insulated box, along with the refrigeration unit and preamplifier was loaned by Dr. D.M. Hunten of the Physics Department, U. of S. The detector lens is not the proper one to make an image of the plates fill the cathode, since its focal length is too

short, but it does not lose any light. The detector seems to be very satisfactory.

The preamplifier is essentially that described by Hunten (1953). The output stage is a 6SN7 twin-triode connected in parallel as a cathode follower. There is a time constant feeding the 6SN7 to reduce noise and an attenuator to feed the Varian Recorder.

The Varian G-10 chart Recorder has a sensitivity of 10 millivolts full scale, and two chart speeds, 2 inches/~~second~~ minute and 12 inches/~~second~~ minute. The output stage, the time constant, zero adjust and output attenuator are mounted on the control panel. The 300 plus and minus is supplied by a Philbrick R-100B D.C. power supply and the negative high voltage by a John Fluke 402 M high voltage supply that is variable from 500 to 1,500 volts.

3.7 General Mounting

The apparatus stands vertical because the plates are horizontal. The North penthouse being only 76 inches from floor to roof is almost ideal as the overall height of the apparatus is 65 inches so this leaves room for one inch at the bottom and ten inches at the top. (See Figure 3.12). This allows one to get his head above the apparatus to adjust the etalon and yet is not so far from the roof that it requires a long light shield and large hole to bring in the light. The hole in the roof is 8 inches in diameter and is fitted with a galvanized chimney and cover.

As mentioned before, everything is mounted from the base

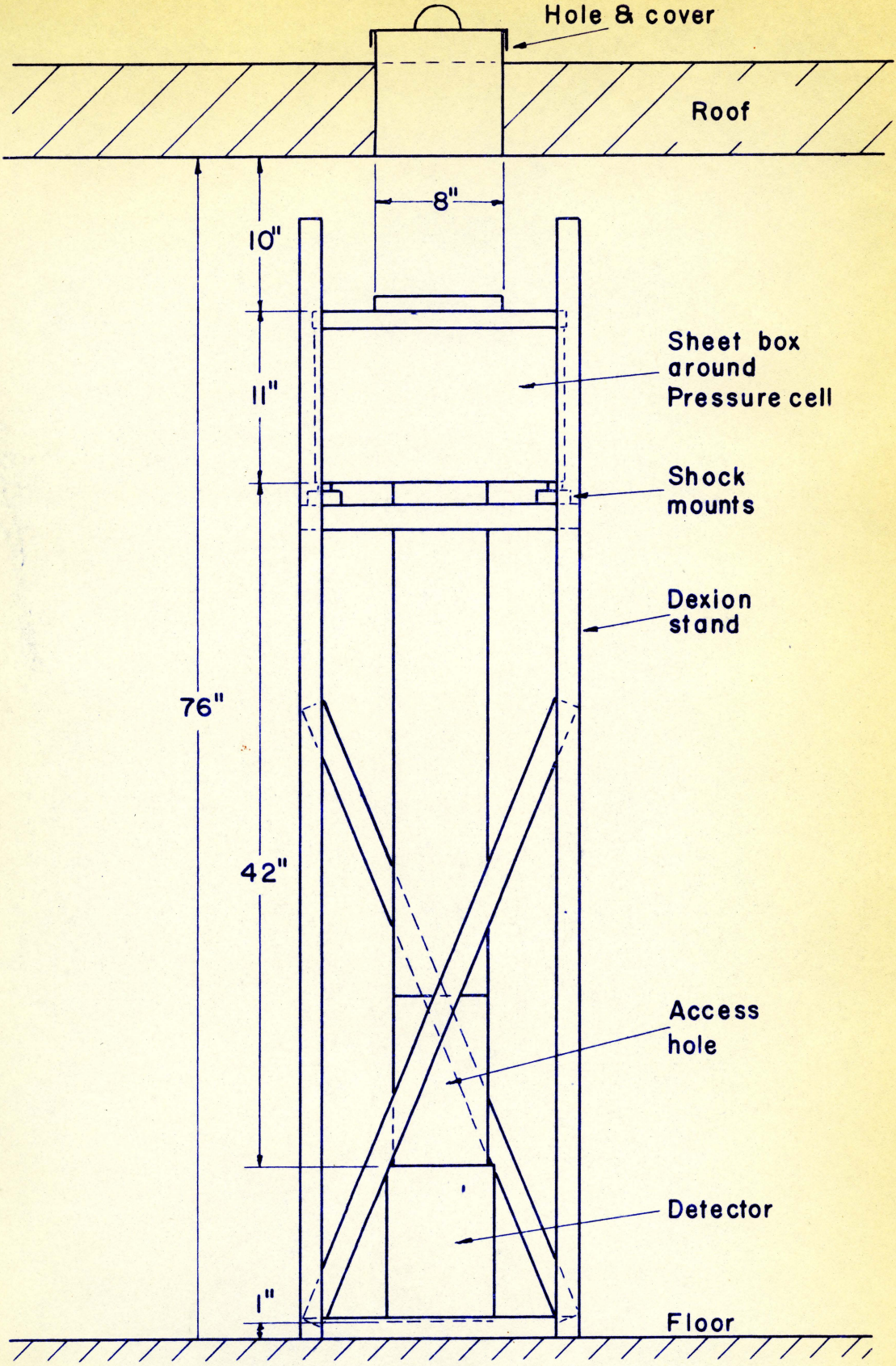

Hole & cover
Roof
8"
10"
11"
Sheet box around Pressure cell
Shock mounts
Dexion stand
76"
42"
Access hole
Detector
1"
Floor

FIGURE 3·12a

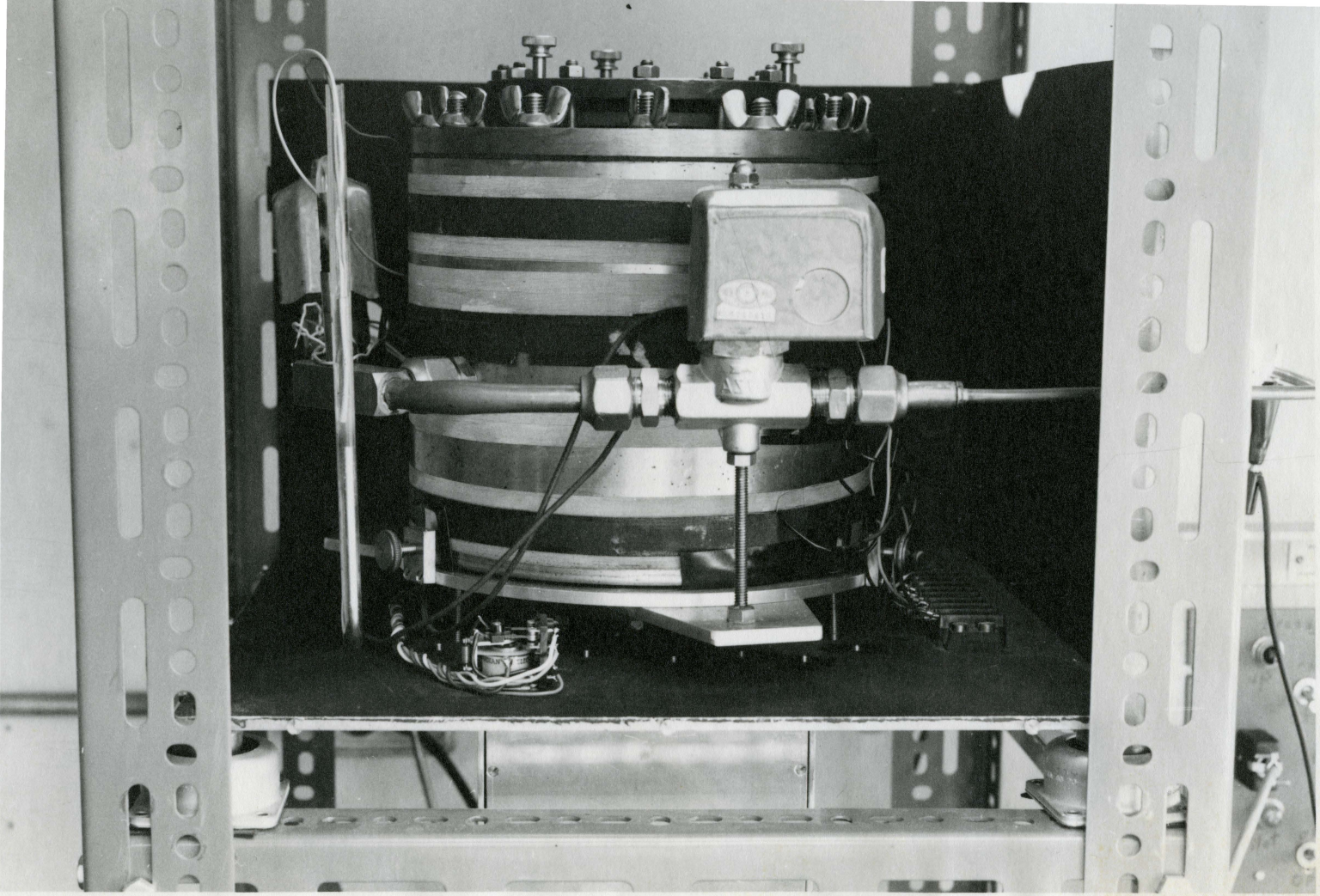


FIGURE 3.12b

plate which is ¼ inch thick aluminum. The square tube that holds the aperture and detector is made of 1/8 inch thick aluminum with ¼ inch brass in the corners, since it must be quite rigid to hold the aperture steady with respect to the etalon. The box surrounding the pressure cell is made from aluminum sheet, and the top and two sides are easily removable.

The idea behind the mounting was to have the whole instrument as separate from the building as possible to cut out vibrations which can be seen when transmitted to the etalon. This is especially important because of the board floor of the penthouse which is not as solid as the rest of the building. The whole apparatus then is hung from the base plate which in turn rests on four damped coil springs (Barry Mounts) which are resting on the dexion cross bars. Thus the instrument does not touch the building except through the springs and the copper tubes that are for the pressure system, but as these tubes are ¼ inch and quite long they are not capable of transmitting much force. Dexion was used for the supporting stand because of its ease of handling and its strength. It does not need to be solid as the apparatus is on springs anyway. A very useful thing about dexion is that if a new detector of a different length were put on, the cross bars could be easily raised or lowered thus making it easy for the apparatus to accomodate it. A mirror with geared sights on it was made to look at any aurora up to about 45^{o}, which is set on

the roof when observing.

The control and power supplies are rack mounted close to the instrument with two eight wire cables leading to it, one for the heating and pressure controls and the other for the preamplifier. The Varian Recorder is on a shelf of the rack mounting at a convenient height. The refrigerator is behind the apparatus close to the detector.

IV THE REAL ETALON

4.1 Introduction

It was noted in the theory that the recorded function Y can be broken down into the source function B, and the instrumental profile W. Thus to get B from Y, W must be known, and since $W = A \not\ast D \not\ast F$ it can be studied by studying A, D and F. The detector function F is rectangular and easily determined so that leaves A, the Airy function and D the defect function. The Airy function and defect function make up the etalon function E.

The method used to study E and D was that described by Chabbal (1958). The instrument is operated at low order;[1] that is t is small and hence $\Delta\sigma$ is large. This is done to make all half-widths in $Y = B \not\ast A \not\ast D \not\ast F$ negligible compared to A and D. The half-widths may be estimated as follows. For a spacing of t = 0.3 mm, $\Delta\sigma = 1/2t = 16.67$ cm^{-1}. For nine-layer coatings $N_R = 180$ and so a = 92.7 mk. Using a 1/8 inch aperture f = 20.2 mk and for the krypton b = 50 mk. Thus if $N_D = 30$, d = 555 mk which is much larger than all the other half-widths, which are negligible except for a. Two methods were used to study the etalon function and defect function and they are

(1) The 0.3 mm quartz spacers were made by E.C. Turgeon who also gave assistance while taking results at low order.

discussed in Sections 4.3 and 4.4.

4.2 Defect Function

The defect function D is defined by $dS = D\, d\sigma'$ where dS is the effective area at the spacing t' such that σ' satisfies the equation $m_o = 2\, n\, t' \sigma'$ as discussed in Chapter II. This can be derived for a simple case so perhaps a better insight into D can be obtained if it is done here.

The case of curvature of the plates is quite common so for simplicity it will be assumed that one plate is spherically curved and the other is flat. Referring to Figure 4.1, dS is the surface element of revolution given by

$$dS = 2\pi R\, dt'$$

Differentiating $m_o = 2\, n\, t' \sigma'$

gives $$d\sigma' = \frac{\sigma}{t'} dt'$$

which to a very good approximation is

$$d\sigma' = \frac{\sigma}{t} dt'$$

$D = \frac{dS}{d\sigma'}$ so by substitution $D = \frac{2\pi R t\, dt'}{\sigma\, dt'}$

Integrating over the whole surface gives

$$S = \int_{s=0}^{s} dS = \int_{t-\frac{\delta t}{2}}^{t+\frac{\delta t}{2}} 2\pi R\, dt' = 2\pi R\, \delta t$$

Substituting for $2\pi R$ in D gives

$$D = \frac{t s}{\sigma \delta t} \quad \text{for } t - \frac{\delta t}{2} \leq t' \leq t + \frac{\delta t}{2}$$

D is thus a constant over the range δt and hence is a rectangular function. The half-width in terms of σ is $d = \frac{\sigma}{t} \delta t$. Hence D is given by $D = \frac{t s}{\sigma \delta t}$ for $\sigma_i - \frac{\sigma \delta t}{t\ 2} \leq \sigma' \leq \sigma_i + \frac{\sigma \delta t}{t\ 2}$ and zero

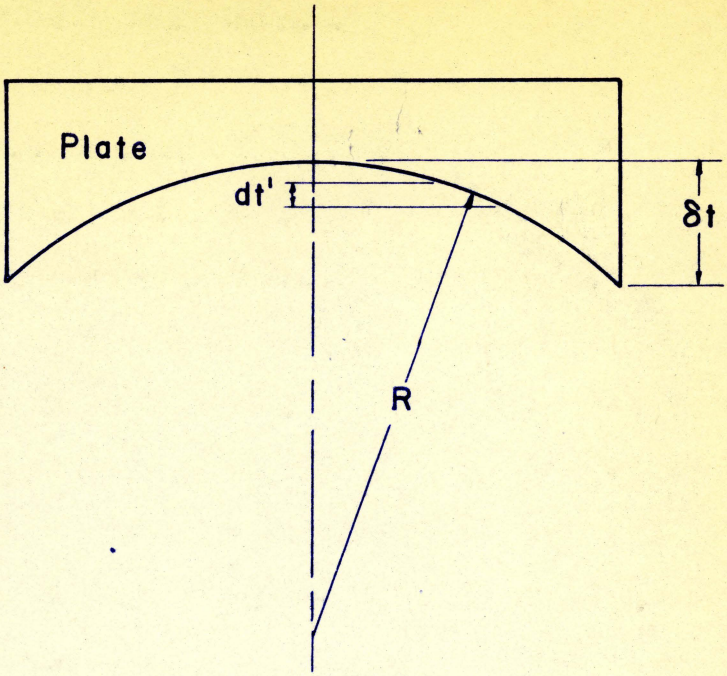


FIGURE 4·1

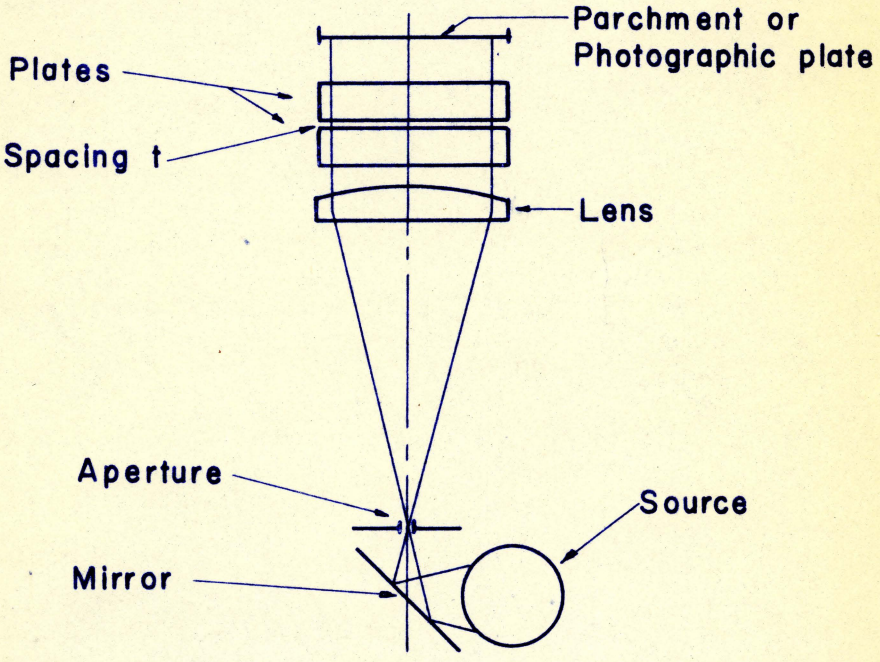


FIGURE 4·2

everywhere else. It is easily seen that $S = \int_{\sigma'=0}^{\infty} D(\sigma')\, d\sigma'$.

The other common defect results from the roughness left after polishing. It is assumed that any curvature is small compared to the roughness. A mean value of T is chosen and it is assumed that the roughness is equivalent to having a large number of etalons of spacing t' and elemental area dS with the spacing t' randomly distributed about the mean value t. A value of σ' is then associated with each spacing t' by using $m_o = 2\, n\, t'\, \sigma'$. The defect function D defined always by $dS = D(\sigma')\, d\sigma'$ is Gaussian in shape as might be expected. In the real etalon the defect function is a combination of curvatures (not necessarily spherical) and roughness.

The plates thus have a topography which can be described in part by contour lines. These contours are lines of constant spacing t and can be found in the following manner. Collimated monochromatic light is passed through the plates as shown in Figure 4.2. This is simply done if the light source is placed below the aperture since the objective lens will collimate the light. The angle the light makes with the axis of the plates is essentially zero and the wavenumber is known, hence the formula $m_o = 2\, n\, t'\, \sigma'$ can be used. The only light that passes through the plates is in the areas where the equation is satisfied, and hence for only a certain value of t'. The light coming through the plates is allowed to fall on parchment paper or a photographic

plate which allows one to observe which areas are transmitting. The line between light and dark will denote a value of t and hence be a contour. In order to get more contours the refractive index is changed so that a different t is required to satisfy the equation.

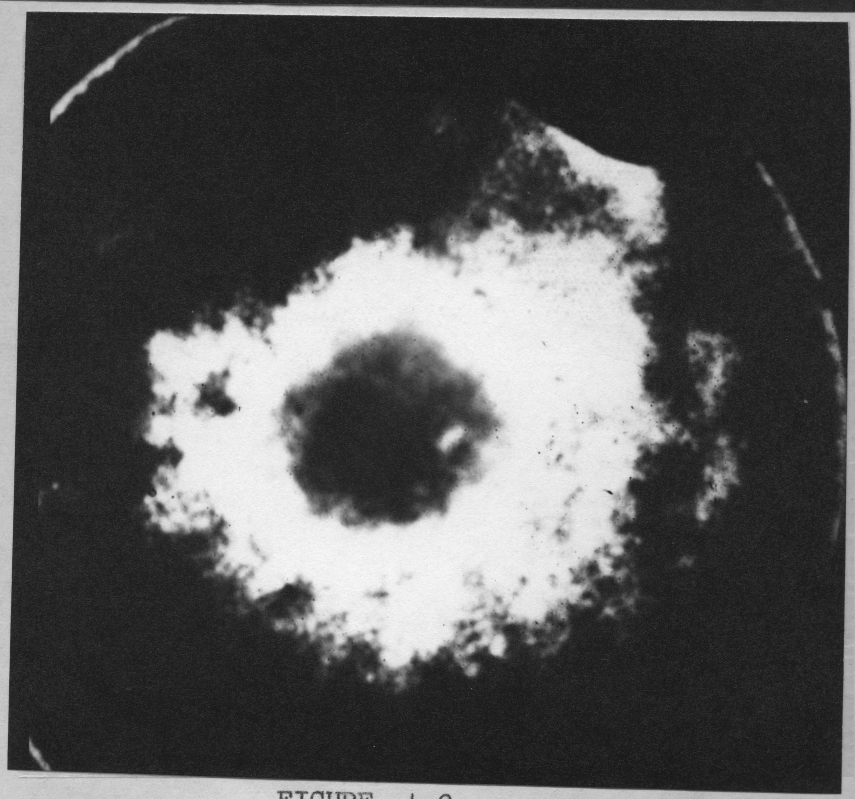
Four photographs were taken to observe the contours and are shown in Figure 4.3. The outside edge between light and dark was drawn for each photograph in Figure 4.4 and the contour heights compared to the first are given in Table 4.1. The minus signs indicate that the plates are concave.

TABLE 4.1

Curve Number	Contour Height (Order/100)	Contour Height (Angstroms)
1	0	0
2	-2.06	-60.6
3	-4.40	-130
4	-7.35	-216

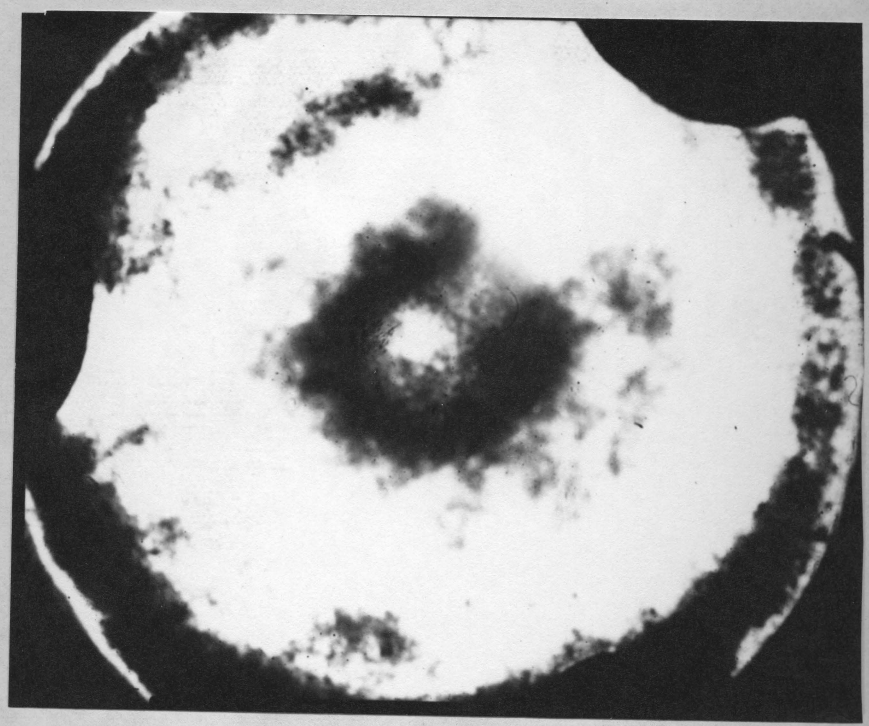
If everything (a, b and f) was of negligible half-width compared to d, and if there was no roughness, the light area would only be a fine line showing the contour. However, since the other half-widths are not negligible, especially a, and there are polishing defects there is a broad area that transmits

#1

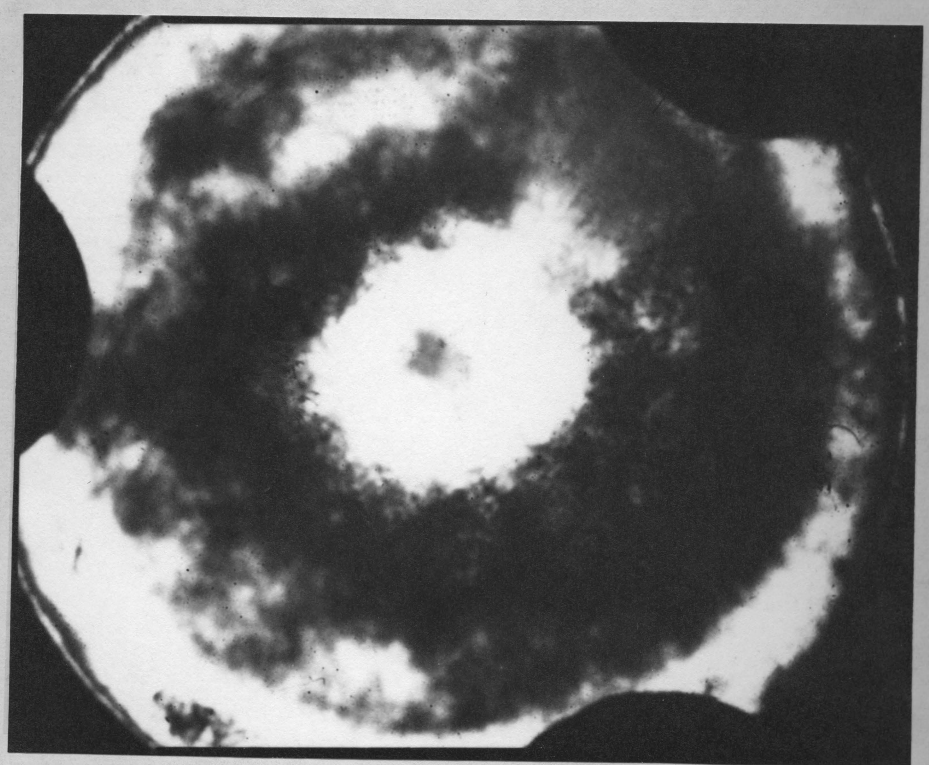


#2

FIGURE 4.3



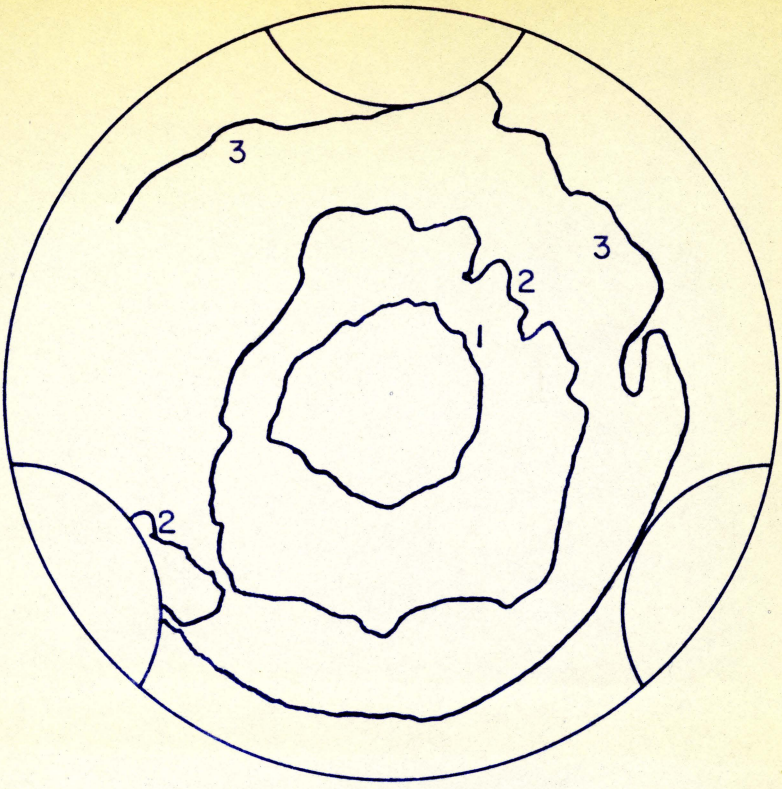
#3



#4

FIGURE 4.3

FIGURE 4·4



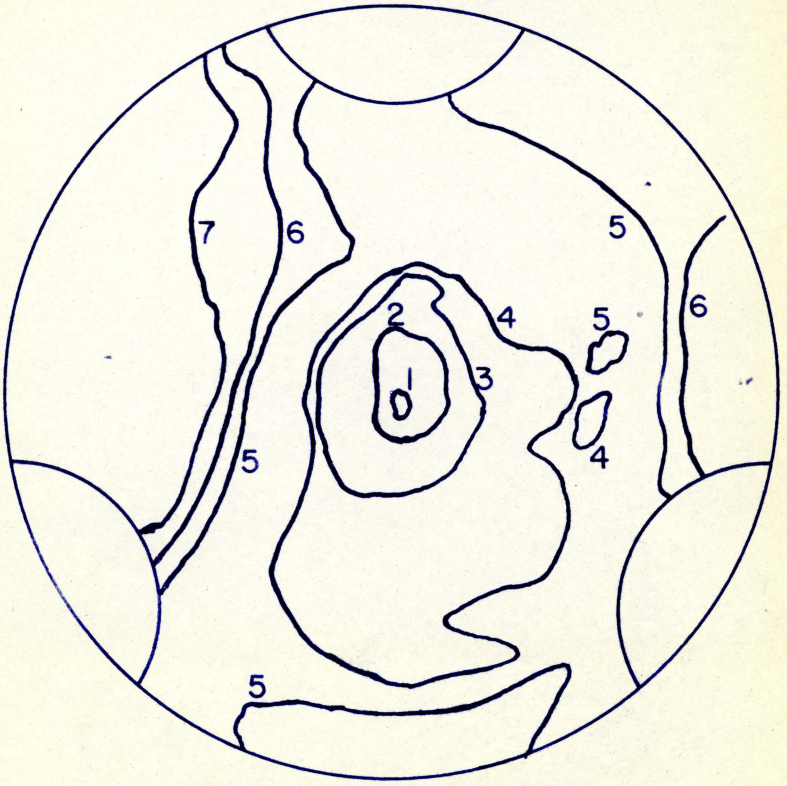


FIGURE 4·5

light. The mottled effect on this area is due to the larger polishing defects. Sodium light was used to observe the contours because of the high intensity required. This accounts for the second light area appearing in the fourth photograph since the two sodium lines happen to be close together at this spacing, t = 0.3mm. The ¼ inch diameter aperture was used also because of the intensity of light required. If a fine line like krypton 5570 A and the 1/8 inch aperture were used more detail would be visible.

4.3 The Etalon Function

The method used to observe the etalon function is to operate the instrument at low order. A spacing of t = 0.3 mm was used and a 9 layer dielectric of zinc sulfide and cryolite was used on the plates. As shown in the introduction, this gives for $\Delta\sigma = 16.67\ cm^{-1}$, a = 92.7 mk, krypton b < 50 mk and f = 20.2 mk. The graph shown in Figure 4.6 is a typical example of Y the recorded function when under these conditions. Utilizing Graphs V 4 and V 5 from Chabbal's (1953) paper the contribution of the various half-widths to y can be ascertained. The finesse of the recorded function Y is 33 so since $\Delta\sigma = 16.67\ cm^{-1}$, y = 505 mk. Y is almost exactly a Gaussian function so this can be used when calculating half-widths. Using the graphs mentioned above and equation $y = b \oplus w$, w becomes 503 mk as compared with 505 mk for y. The small effect of b (50 mk) is easily seen. The detector

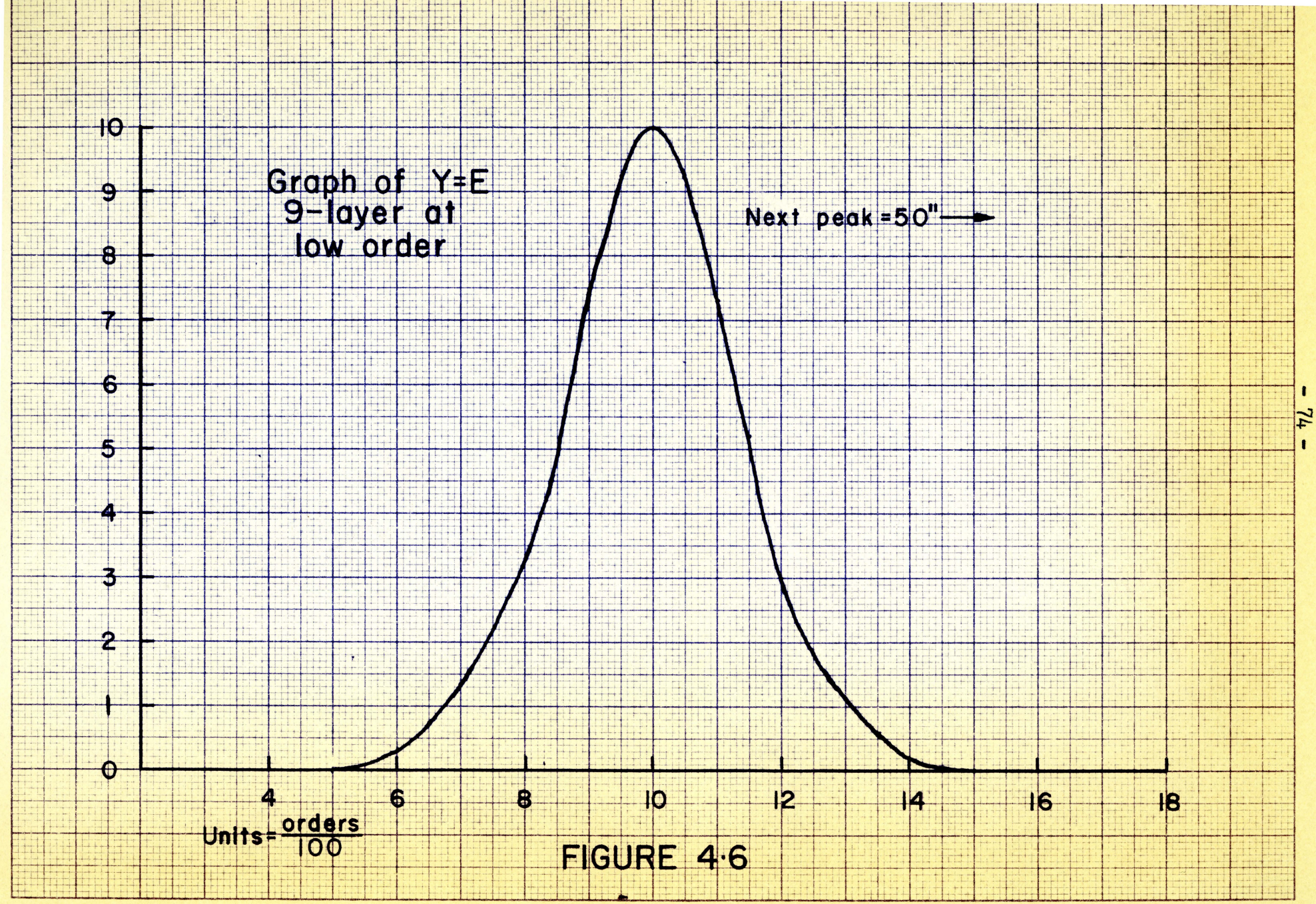


FIGURE 4·6

function F is so small (f = 20.2 mk) that it can be neglected. Thus since B and F in Y = B ∗ E ∗ F can be neglected, it can be said to a good approximation that Y ≡ E. d can be obtained from a⊕d = e assuming a = 92.5 mk and becomes 447 mk, or N_D = 37.4. The etalon function is thus easily obtained by simply operating at low order, and running the instrument over a large wavenumber interval.

4.4 The Effects of Adjustment

All of the forgoing has assumed that the plates are in adjustment, without saying what best adjustment is. The adjustment is closly linked with the defect function. For example, if the plates are perfectly flat and parallel the defect function is a Dirac delta function with area equal to that of the plates. If however, the plates are put out of adjustment then the defect function becomes a rectangular function of the same area as before. The adjustment of the plates greatly affects the etalon function. A criterion for good adjustment could then be to make the etalon function as narrow as possible. The effect of bad adjustment can be seen by refering to Figures 4.5 and 4.7. Figure 4.5 shows the contours drawn on parchment when the plates are not as well adjusted as when the photographs were taken for Figure 4.3. The contour heights are given in Table 4.2. There may have been other effects such as poor temperature equilibrium since the etalon must be in the temperature controlled cell for

a day to ensure good equilibrium.

TABLE 4.2

Curve Number	Contour Height (Order/100)	Contour Height (Angstroms)
1	0	0
2	- 2.35	- 69
3	- 4.12	-121
4	- 5.59	-164
5	- 7.05	-208
6	- 8.52	-251
7	-10.0	-294

It can be seen from the contours that if the left hand side were raised the contours would become more circular. Thus, the way in which to obtain the best adjustment is to make the illuminated area on the parchment the greatest when observing the defects. If a polythene film is used instead of parchment, both the Fabry-Perot rings and the contours may be seen which is a great aid in adjusting the plates. At high order the defects cannot be easily seen so the method used is to keep the Fabry-Perot rings the same diameter for moving the eye over the greatest area of the plates, rather than just at the spacers.

For Figure 4.7, which is the output Y, the plates were

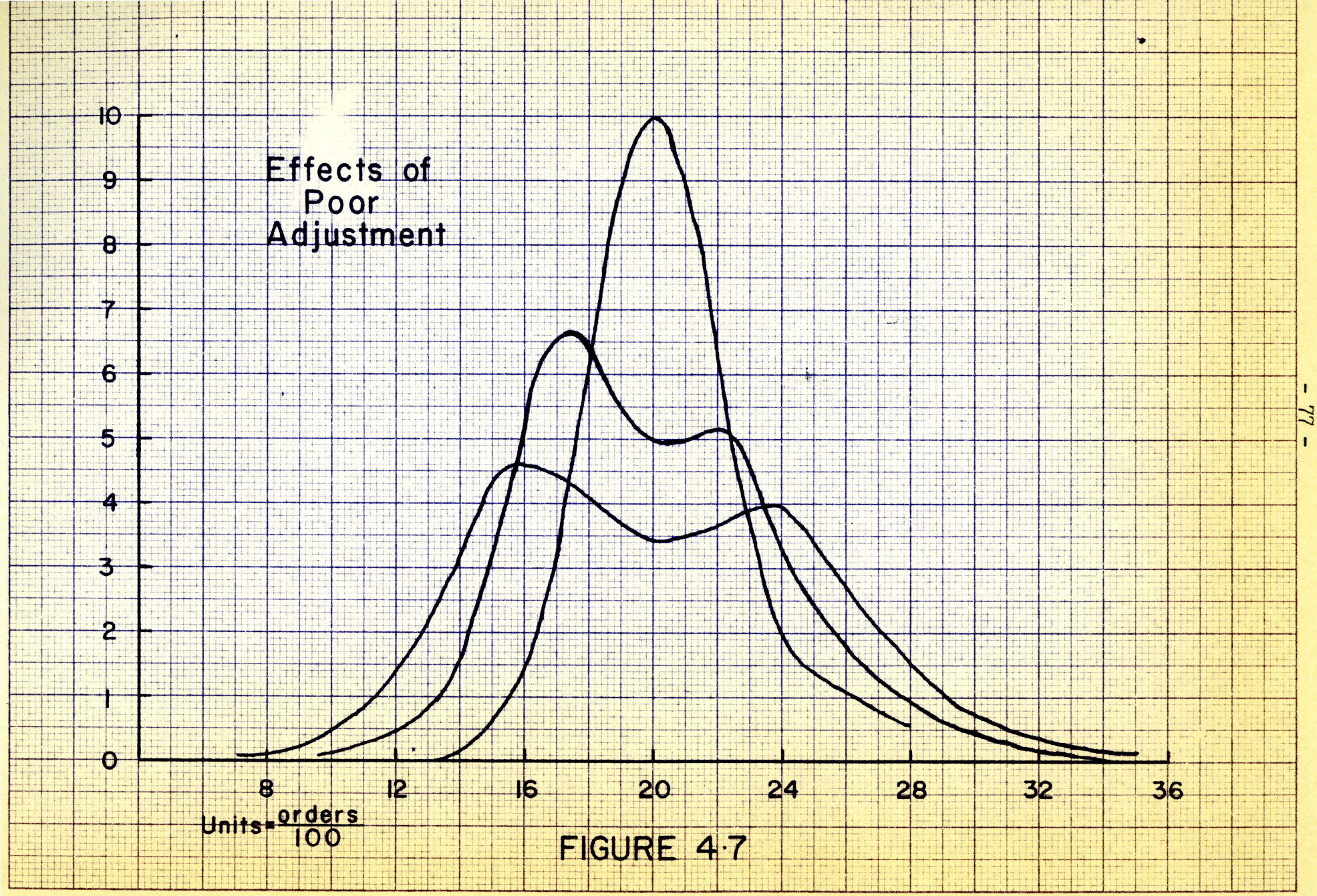


FIGURE 4·7

purposefully put out of adjustment to see what would happen to the etalon function. This was done when 7 layers were on the plates, so the central curve is when the plates are in adjustment. The next highest curve was made when one spacer was compressed 0.042 orders. This was done by checking the diameter of the Fabry-Perot rings while in adjustment, changing the refractive index the required amount and bringing only one spacer the same as it was before. A similar procedure was followed for the lowest curve, the one spacer being compressed 0.091 orders.

The results show that when put out of adjustment the etalon function tends to become rectangular which gives a defect similar to curvature (Chabbal (1953)). It is also seen that there is no reason to expect the etalon function to be exactly symmetrical. The peak of the curve is of course lowered because the defects just change the spectral distribution, not the total light flux in one order. The two peaks cannot be explained, as unfortunately the instrument was not put out of adjustment when using nine layers where the defects are easily observable. However, it appears that when put out of adjustment the plates have the equivalent of two major areas slightly separated from each other so that first one transmits and then the other.

4.5 The Effects of Masking

It can be seen from the contours of Figures 4.3 and 4.4 that the plates are concave. It should then be possible to

narrow the etalon function by using a smaller area of the plates, since the defect function will be smaller. One day the output function appeared as shown by the outside curve of Figure 4.8. The bump on the front of the curve must be made by defects , since the only other function (Airy) is symmetrical. It should, by masking the correct part of the plates, be possible to make the bump disappear. The central one inch diameter of the plates was masked and the large solid curve of Figure 4.9 resulted, which has no bump on the leading edge. Next the plates were masked except for a one inch diameter hole at the center and the low solid curve of Figure 4.9 resulted. When these two are added together they should result in the curve obtained in Figure 4.8 with no masking. The dotted curve in Figure 4.9 is the result, not exactly the same as in Figure 4.8, but nearly so. It may be concluded that the hollow in the center of the plates caused the lump on the front of the recorded curve. The same can be done to the back of the curve by masking out the outside edge of the plates. This was done and the back edge of the etalon function was reduced as would be expected.

If a small portion of the plates is used the defects should be caused by the roughness left after polishing, since the curvature should be small. A good portion of the plates was selected and a one half inch diameter hole left, the rest of the plates being masked. The resulting curve , shown as the narrow one in

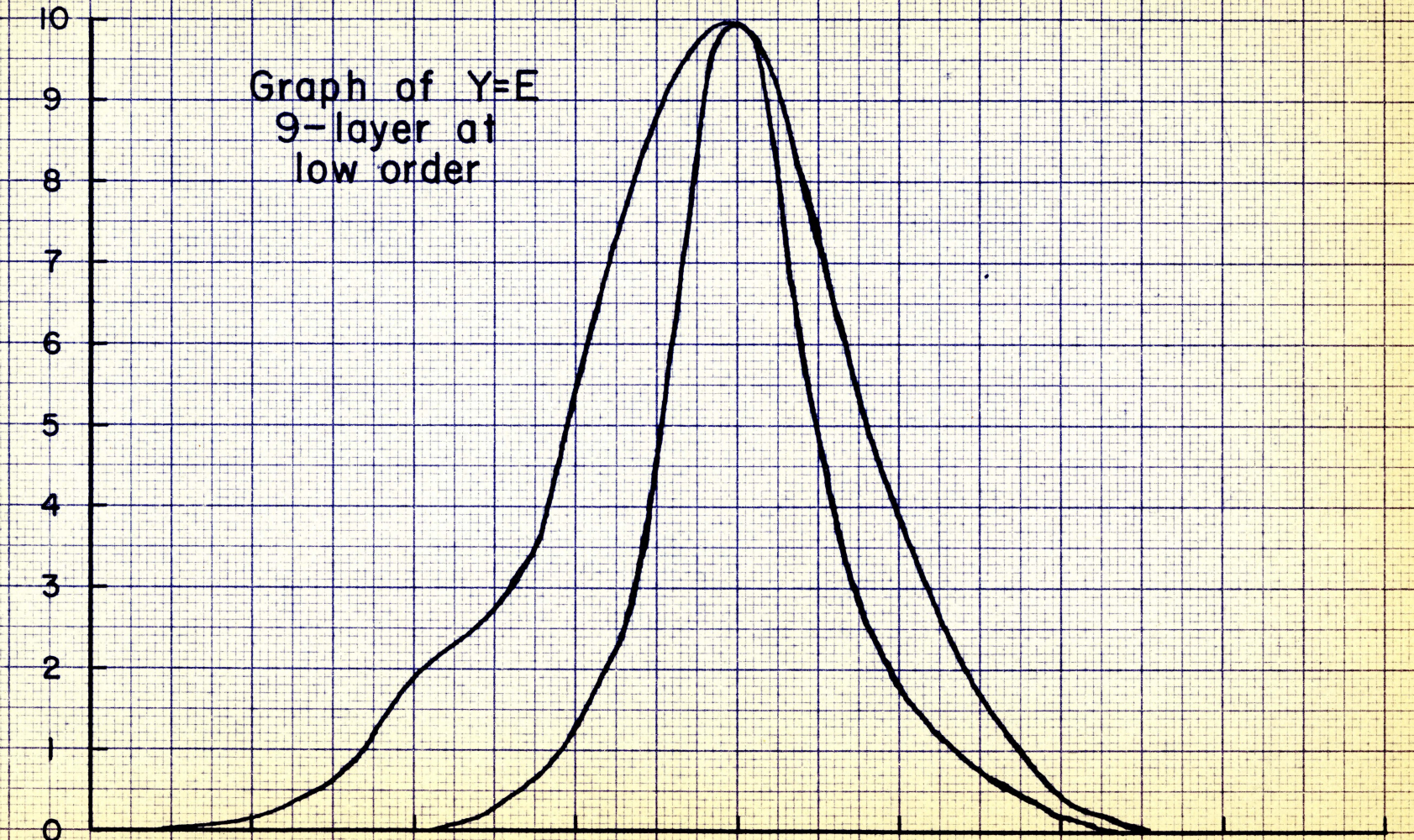


FIGURE 4·8

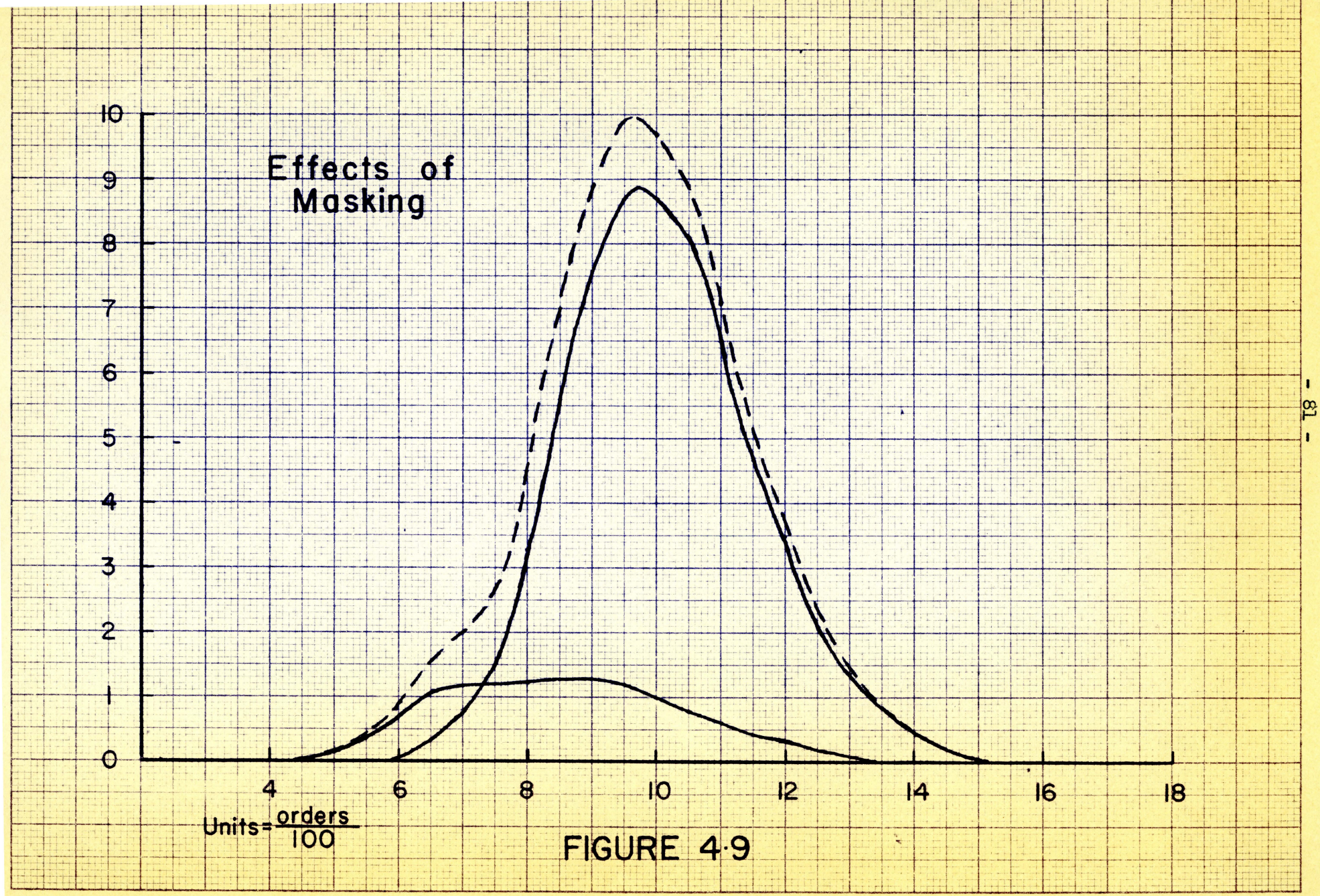


FIGURE 4·9

Figure 4.8, has a finesse $N_E = 53$. Since $t = 0.3$ mm was used $y = e = 315$ mk. The Airy function can be taken off this in the manner used before by using $e = a \oplus d$, and this leaves the defect function with a half-width of 254 mk and $N_D = 65.5$. Thus the defects due to polishing are much less than the defects due to curvature.

4.6 Maximizing L X $\underline{R}$

It was stated in the theory that the product of luminosity and resolution can be maximized by using the proper reflectivity to suit the value of the defect function. As dielectric multi-layers are being used there are only step-wise changes in N_R allowed. Using zinc sulfide and cryolite as the dielectric the finesses are given by Chabbal (1953) as 9 layer - 180, 7 layer - 60 and 5 layer - 23. The nine layer coatings on the plates, which were used for measuring defects and the first results are obviously too highly reflecting, since the defects are at best around $N_D = 35$. For good L X $\underline{R}$ one should have $N_R \doteq N_D$. The only spacer available at the present time is a 1 cm spacer for high order, so it seemed reasonable to pick the reflectivity that would still give the required resolution at this spacing. It turns out that seven layers are satisfactory at a spacing of 1 cm and better at 2 cm or 3 cm of course. However five layers are not satisfactory until a 3 cm spacer is used and there may be danger from overlapping at this spacing at this lower finesse.

The calculations were made assuming the auroral green line has a half-width of b = 60 mk and the ratio of b/y = 0.8 and b/y = 0.9. The ratio b/y gives an idea of how closely the recorded function will resemble the source. As b/y approaches 1, the recorded function approaches the source function and the luminosity received by the photomultiplier decreases. It was thought that b/y = 0.8 was satisfactory and b/y = 0.9 would be valuable if it could be attained. With seven layers it is possible to get b/y = 0.8 with 1 cm spacer and b/y = 0.9 with a 2 cm spacer.

The plates were vacuum evaporated with seven layers of zinc sulfide and magnesium fluoride. This combination was used in place of zinc sulfide and cryolite because it is found when making interference filters the passband is larger when using magnesium fluoride in place of cryolite, which indicates a smaller reflectivity. The resulting etalon function has half-width 816 mk and finesse of 20.2 and a tracing done at low order is shown in Figure 4.10. A Gaussian was plotted with the same half-width and found to fit exceptionally well except for the foot of the curve on the left side. It was not expected to be as Gaussian as the nine layer etalon function since the Airy function has a larger role. Assuming that the defects are Gaussian and combined with an Airy function the value of a and N_R can be found. Since $a \oplus d = e$ and e = 816 mk, and assuming N_D = 30 and 35, the values of a and N_R become a = 456 mk,

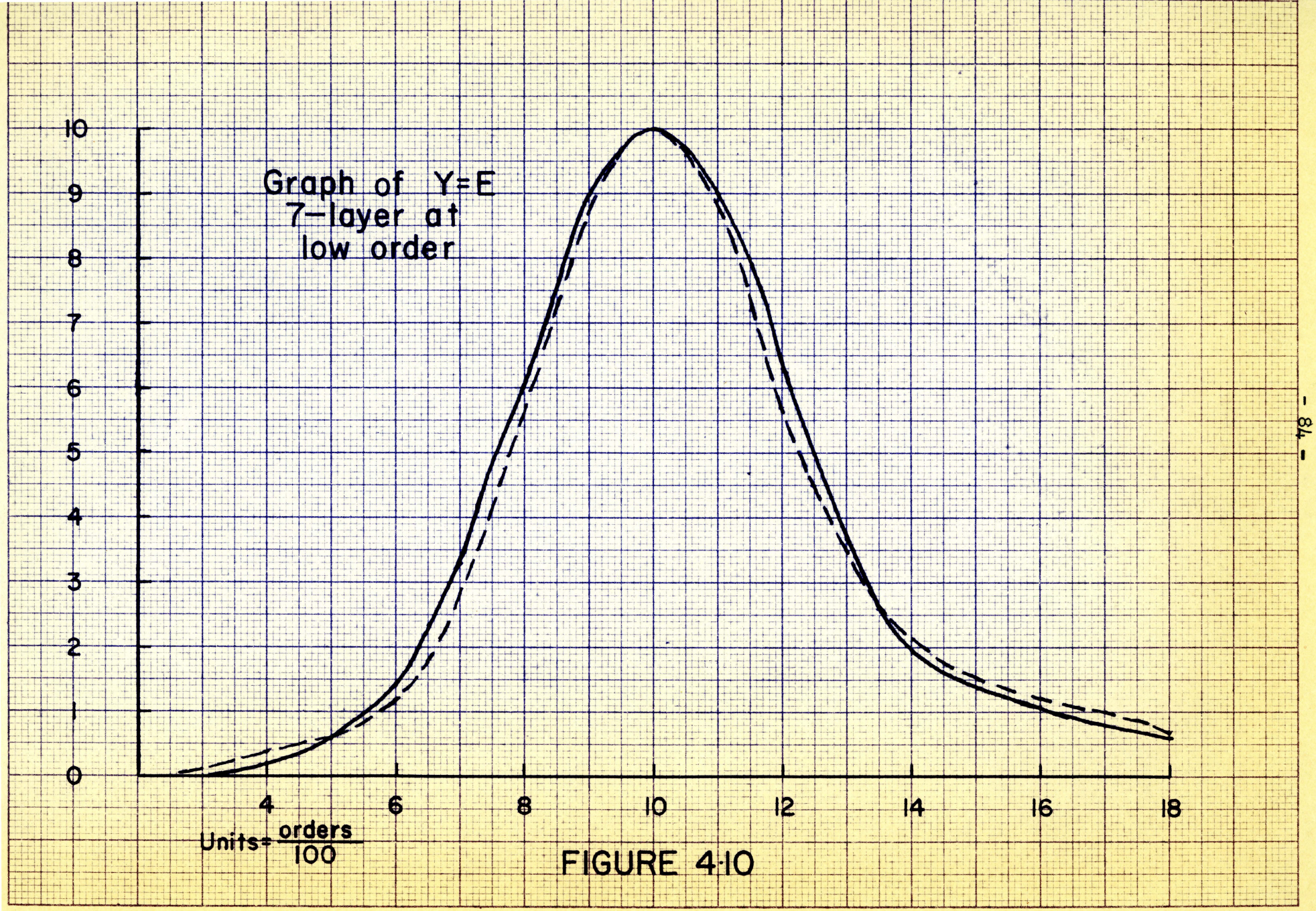


FIGURE 4-10

$N_R = 36.5$ for $N_D = 30$; and a = 520, $N_R = 32$ if $N_D = 35$. It would have been valuable if R could have been measured indepently for both the seven and nine layers to see how closely they come to the results given by Chabbal (1953). It thus appears that the value of N_R is about correct now, but since it is lower than $N_R = 60$ for seven layers of zinc sulfide and cryolite it appears that perhaps t = 2 cm, for high order, would be better.

The change in the Airy function when changing from nine to seven layers can be further illustrated. The inside dotted curve in Figure 4.10 is the recorded function when the plates are masked except for the central one inch. It is seen that this curve is nearly the same size as that for the whole etalon which is in contrast to the two curves shown in Figure 4.8 for nine layers. This means then that the Airy function is broad enough that very little gain in resolution is made by decreasing the area of the plates in contrast to the great gain in resolution made by masking at nine layers.

4.7 The Instrumental Profile

The instrumental profile W is of great importance since it must be known to obtain B from Y. In the following sections the etalon function has been studied and this is broadened by F, the detector function, to give W. It was found that the etalon function when in good adjustment was very close to Gaussian for both nine and seven layers on the plates. Thus it is easy to

calculate the value of w from $f \oplus e = w$ utilizing the curves V4 and V5 by Chabbal (1953). If it is assumed for nine layers $N_E = 35$ and the 1 cm spacer is used. e = 14.4 mk and since f = 20.2 mk (1/8 inch aperture), w = 21.9. For seven layers if $N_E = 20$ and 1 cm spacer is used then e = 25.2 mk and f = 20.2, w = 30.1. These and other values will be used in Chapter V.

The instrumental profile W can also be obtained if a known spectrum is observed. A narrow symetrical line such as the Mercury 198 green line is very useful for this since its shape is known. $Y = B * W$ may be used as Y and B are known; thus W can be found by Fourier transforms. The half-width w from $y = b \oplus w$ can also be obtained. This method is very good since the instrument does not have to be altered in any way from observing the unknown spectrum to the known spectrum, except for perhaps changing a filter.

The fact that the etalon function is very dependent upon the defects, including adjustment, is very important. This means of course that the instrumental profile is also dependent upon the defects and it may change in shape and even become asymmetrical as was seen when the etalon was not well adjusted.

The result of this variability in the etalon function is that the resolution should be higher than would be necessary if the instrument had a constant stable etalon function. This can be overcome by making periodic records of the known line during

an evenings observation. However, if the instrumental half-width is narrow enough compared to the source it does not matter, within a certain limit of error, whether the instrumental profile is rectangular, Gaussian, Airy or even a bit asymmetrical.

V RESULTS

5.1 Introduction

Most of the results were taken before the plates had been observed at low order. Thus there was no idea of what the instrumental profile would be. A standard source Kr 5570A was used to obtain the instrumental profile. There were good results obtained on May 31st, June 3rd and June 5th, 1960 with nine layers on the plates and some more results using seven layers on July 12th and 13th, 1960.

At the time of observation with nine layers on the plates, there was little hope of getting any results from aurora because of the low transmission. But since at that time there happened to be aurora the instrument was tried. When results were obtained it was suprising, even though auroral intensity of 1 to 2 was required. The results also showed that the plates were satisfactory even before they had been observed at low order.

In Chapter IV it is seen that the etalon function appeared to be Gaussian, for both the nine and seven layer results. This fact was used for the calculations to follow. The proper method to use is to take Fourier transforms of the recorded function, when using the standard source, and the transform of the standard source function to obtain the etalon function. This etalon function is then used on the auroral recorded function to obtain the auroral profile. This method was not possible as the source

function Kr 5570 A was not known. The method used, then, was to consider only half-widths of the important functions.

A value of the half-width of the krypton line was found by using the etalon function from Chapter IV. The assumptions made were that the best value of finesse which the etalon can have is 35 (e = 14.4 mk) for nine layers and 20 (e = 25.2 mk) for seven layers, and that a record of the krypton was obtained with these values of N_E. This record was found by picking the best value of N_Y (the finesse of the recorded function Y) for the krypton from all those obtained.

From $b \oplus e \oplus f = y$, and using the curves V 4 and V 5 given by Chabbal (1953), a value of $b \oplus e$ can be found from this best value for y, since f is known and $b \oplus e$ is Gaussian. Using $b^2 + e^2 = (b \oplus e)^2$, which is valid for Gaussian functions, and the value of e given above, b_k (for krypton) can be found. For nine layers the best result for N_{Yk} of krypton was 11.9 giving $y_k = 42.7$ mk. This gives a value for b_k, the half-width of krypton, of 37.5 mk. In the seven layer results the best value of N_{Yk} was 11.0 with $y_k = 45.8$ mk. The result for seven layers is $b_k = 35.3$ mk. These two values of b_k were averaged and the value $b_k = 36.0$ mk was used for the following calculations.

There is now enough information to calculate a temperature for auroral records. The half-width of the etalon function, e, is required and can be found if a scan has been obtained using the krypton source. Usually the krypton recorded function y_k does

not change much from one set to the next so the mean value of N_{Yk} for these two sets can be used for all the aurora between them. From this N_{Yk}, y_k is found which gives $y_k = b_k \oplus e \oplus f$. Since b_k and f are known, e can be found. It is then assumed that this is the same e that is operating on the auroral profile. The finesse of the auroral recorded function N_{Ya} and y_a are measured and the effect of f removed to give $b_a \oplus e$. Using the e obtained above, b_a is known (since $b^2 + e^2 = (b \oplus e)^2$ for Gaussian functions). Once b is found a temperature can be easily calculated from $b = 7.16 \times 10^{-7} \sigma \sqrt{\frac{T}{M}}$ which simplifies to $T = 9.76 \times 10^{-2} b_a^2$, where b_a is given in millikaisers, for the auroral green line.

5.2 Observations

When observing, the instrument was allowed to scan over three or more orders per pressure change, or cycle, and the observation position was not changed during this time. If the curves looked similar it was assumed that the fluctuations in intensity were not rapid enough to be harmful and these results were used. The finesses for the similar curves were averaged together and the result used as N_Y. A list of results is given in Table 5.1 in order of increasing temperature. A value for N_{Yk} that was used for the krypton and N_{Ya} for the aurora is given so that the temperatures can be recalculated if a better value of the half-width of krypton is found at a later date.

TABLE 5.1

	T °K	Auroral Type	Number of Layers	N_{Y_k}	N_{Y_a}
*	222	arc	9	7.20	6.45
*	283	arc	9	9.26	7.45
	288	arc	9	9.00	7.28
	314	arc	9	9.00	7.07
	342	arc	9	9.26	7.00
*	344	diffuse	7	11.00	7.66
	346	arc	9	9.00	6.95
	348	arc	9	9.00	6.94
	352	arc	9	7.20	5.84
	366	glow	9	7.20	5.77
	396	arc	9	9.00	6.54
	406	arc	9	8.27	6.21
	408	diffuse arc	7	11.00	7.10

continued next page

TABLE 5.1
(continued)

	T °K	Auroral Type	Number of Layers	N_{Y_k}	N_{Y_a}
	420	glow	7	7.20	5.57
	410	arc	9	8.91	6.45
*	422	diffuse	7	11.00	7.10
	424	rays & glow	7	9.96	6.74
	441	glow & mixture	9	8.27	6.08
	450	arc	9	9.00	6.81
	452	band	9	8.27	6.00
	465	arc or glow	7	9.96	6.47
	474	arc	9	8.91	6.10
	487	glow	7	9.96	6.39
	541	glow	7	9.96	6.06
*	570	glow	7	9.96	5.96
*	702	glow & mixture	9	8.27	5.14

* see Table 5.2

From Table 5.1 the large temperature range covered can be seen. It is not easy to assign a limit of error to the values but a maximum possible and minimum possible temperature can be assigned in the following manner. Obviously, if it is assumed that $N_E = \infty$, that is, the etalon function $E \equiv 0$ and the instrumental profile W is equal to F, the detector function, then nearly all the recorded line width will be due to the auroral line since F is small. This will give a maximum possible value of b_a and also T. To obtain the minimum possible value it is assumed that the krypton line is temperature broadened only, and is at room temperature. This will make e large and hence b_a and T small. The b_k assumed to be 36 mk gives a temperature of the krypton, if it is pure Doppler broadened, of 663 °K so if a temperature of 300 °K is assumed b_k becomes 24.5 mk. Using these two extreme cases some sample temperatures, marked with a star in Table 5.1, were calculated and shown in Table 5.2.

In the first column the measured values of the half-width of the auroral recorded function are given. In the next column the values of e used in the calculations are listed. The third column contains the values b_a obtained from the combination of $b_a \oplus e$ after the effect of the detector function is removed. In the temperature column the minimum possible, calculated, and maximum possible temperatures are given and in the last column b/y, which is a measure of the adequacy of the resolving power,

TABLE 5.2

	y_a mk	e mk	b_a mk	T °K	b/y
min	78.2	65.6	39.8	154	.510
*		60.0	47.8	222	.615
max		0	76.7	575	.980
min	67.6	46.4	47.0	215	.695
*		38.0	54.0	283	.799
max		0	66.0	425	.976
min	65.7	36.3	52.8	272	.806
*		25.2	58.6	344	.894
max		0	64.1	400	.976
min	72.0	36.3	60.3	354	.839
*		25.2	65.8	422	.915
max		0	70.5	487	.979
min	84.5	41.9	71.7	503	.848
*		32.4	76.5	570	.905
max		0	81.4	674	.964
min	98.2	53.5	80.5	630	.821
*		46.6	84.6	702	.863
max		0	96.8	910	.986

* see Table 5.1

is tabulated.

The range between maximum and minimum values assigned in Table 5.2 vary between 400 °K and 133 °K. This large range is to be expected because of the method used in the calculations.

In Figure 5.1 an exact copy of two traces is shown including noise. They were made within 5 minutes of each other. The solid narrow curves were made from an arc in the south and had a temperature of 406 °K given in the table. The shorter broken curves were made when observing in the east, looking at a mixture of rays and bands and the temperature is 702 °K. The effect of the difference in width would be more obvious if the two sets of curves were normalized to the same height. It would be difficult to normalize the noise so they were shown so that the half-widths would be at about the same place. Each of these sets has a third curve which is similar to the first two and all three curves were averaged in each case to give the value of N_{Ya} for the temperature calculation. It is important that the curves are not overlapping so that the zero level is not raised and it can be easily seen that the orders are far enough apart to make overlapping negligible. The noise level is about average or perhaps greater than average.

When observing with nine layers the chart speed was 12 inches/minute and the distance between orders approximately 5 inches. One order was scanned over in 25 seconds and the important part

Auroral profiles

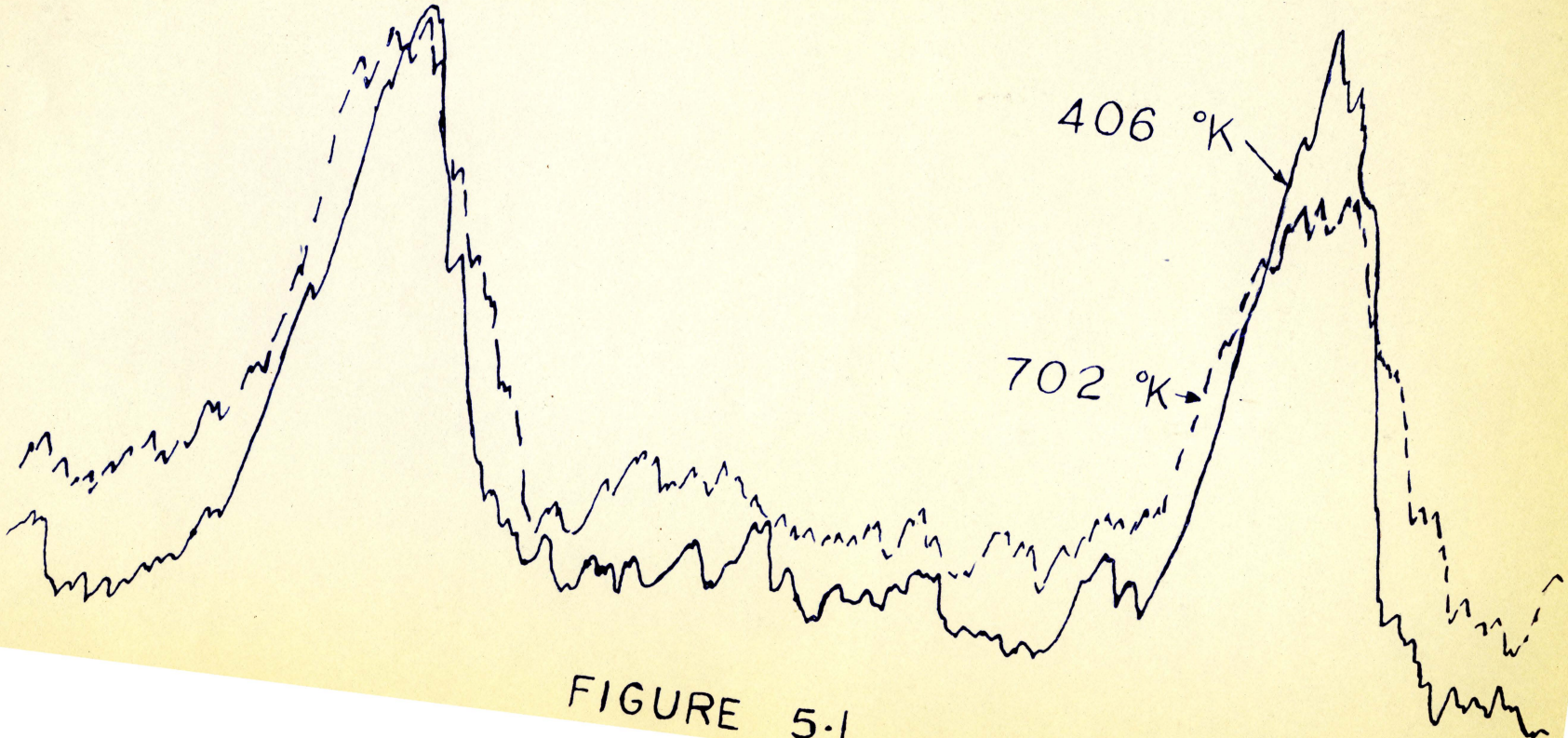


FIGURE 5.1

of the curve in less than ten seconds. The instrument could be described then as rapid-scanning. At seven layers a finer orifice and hence slower scanning speed was used in an attempt to observe night airglow, 5577A. This reduced the scanning time to about 1.6 minutes/order. This may account for the fact that the highest and lowest temperatures were observed with nine layers, since the temperature is not time averaged as much.

The aurora observed with nine layers must be bright because of the large absorption. This aurora is usually associated with low forms which may account for the predominantly low temperatures for nine layers. There were a few results taken on May 31st when it was late in the morning twilight, using a good interference filter to cut out the white light, no stars being visible but the aurora still faintly visible. The temperature of 352 ^{o}K in Table 5.1 is one of these and the temperature is just about average. The majority of aurora seen with seven layers was very faint since the brighter forms appeared to be too fast moving for the slower scanning rate. This may account for the higher temperatures obtained because these forms are generally thought to be of greater altitude.

Recordings were obtained using seven layers when no aurora was visible but they were quite noisy so no temperature could be obtained. There may not have been as low intensity as night airglow however, because there was aurora visible earlier in the

evening. Using a slower scanning rate and long time constant to cut down the noise it should be possible to get usable results on night airglow 5577A.

5.3 Discussion of Results

The results obtained in the previous section range from 222 °K to 702 °K with the average at 411 °K. There are three temperatures between 222 °K and 300 °K, eight temperatures between 300 °K and 400 °K, twelve temperatures between 400 °K and 500 °K, and three temperatures over 500 °K.

These results can be compared to temperatures obtained by other workers with Fabry-Perot interferometers. Wark (1960) obtained no auroral results but for night airglow he found a mean temperature of 184 °K. This can not be compared to any results obtained here, since no airglow was observed. His observations of the oxygen 6300A line showed average temperatures of 710 °K for twilight, 730 °K for aurora and 980 °K for airglow. These may be compared more favourably with the high results obtained here, since the altitude is thought to be higher for forms such as glows and rays. Wark suggests heights of around 210 km for the 6300A emission.

Armstrong (1959) obtained temperatures in the range 180 °K to 235 °K for the night airglow 5577A with the most probable value at 190 °K. For aurora, Armstrong's results compare very favourably with the ones obtained here, being in the range

310 °K to 540 °K. Since more results were obtained here it is not odd that they should cover a larger temperature range. The much greater scanning speed used here may also account for more variation in temperatures.

When the results are compared with rotational temperatures obtained from the 3914A N_2^+ band (Shepherd 1953, and Shepherd and Hunten 1954), they are seen to be much different. In general the temperatures are higher; average of 411 °K compared to a mean of 252 °K for N_2^+ rotational temperatures, and cover a larger range; 222 °K to 702 °K compared to 173 °K to 440 °K. Nevertheless this difference has been shown to be real.

There are some possible explainations for this difference in the two methods. One is that the high temperatures were nearly all from faint forms which may be from higher aurora. There were no faint forms observed by Shepherd.

Another contributing factor to the discrepancy between 3914A and 5577A results could be the larger field of view, 3° in the vertical direction, for the scanning spectrometer used by Shepherd and Hunten, which would tend to average out the temperatures. The field of view for the instrument described here is about 0.17 ° for the total angle at the apex of the cone which includes the field of view.

The distribution of N_2^+ and O are different in the atmosphere. The concentration of O is increasing at auroral heights

and it is believed that the N_2^+ is decreasing, hence the 5577A temperatures may originate from higher altitudes. There is still a slight possibility that the excitation of N_2^+, since it is not forbidden, gives it a temperature which is not the same as the Doppler temperature.

It would be very useful to carry on simultaneous, and closely coordinated, observations for both the 5577A and the 3914A temperatures to check for correlation in form, height, and seasonal variations. This may be possible with the new machine built by E. Rawson at the university here recently since it has a narrow field of view.

These temperatures are intended to be preliminary and when a mercury 198 lamp is used in conjunction with the recording of aurora it should be possible to obtain an actual auroral profile and see whether it is Gaussian or not. However, the results given in Table 5.1 are thought to give real temperatures, probably better than within the broad margin of error assigned in Table 5.2.

5.4 Discussion of the Instrument

From the experience gained in observing, there are some improvements that could be made to make the instrument more useful and more versatile.

If it were possible to choose the scanning rate desired, a fast rate could be used for high intensity fast moving forms,

since a large time constant is not necessary to cut out noise. Then if less intense forms were seen the scanning rate could be decreased and the time constant increased until night airglow could be observed. This could be simply done by having a manifold of perhaps three orifices with valves on the high pressure side which would enable one to choose the proper orifice for the desired scanning rate. The resolution is already variable within limits by changing the aperture which changes F. Thus if very bright forms occur a high scanning rate with corresponding high resolution could be used.

The Varian recorder limits the scanning rate now because of the pen rise-time of 1 second, so if a higher scanning rate than 25 seconds per order is to be used, a different recorder would have to be employed.

It would be interesting to see how fast the pressure could be changed without turbulence becoming important. If it were possible to scan at a fast rate (a second or less per order) an oscilloscope could be used to observe the profiles, and when a good profile was obtained, it could be photographed. The horizontal plates could be driven by a pressure transducer so that the wavenumber scale would be easily obtained.

One thing that limited the value of the results obtained was the fact that the exact point of observation was not known. A mirror system should be built so that it could be easily

pointed anywhere in the sky, and the direction of observation accuratly known. If this were done it could be stated what part of the auroral form was being observed, not just the whole form. If lenses were available to put in front of the instrument, its field of view could be changed to any value desired which may be useful when observing rapidly moving forms that can perhaps be averaged.

It is suggested that a new spacer be made of 2 or perhaps 3 cm spacing. There are two reasons. The first is that if higher resolution can be used on bright forms it would be difficult to obtain this with the one centimeter spacing. The other reason comes about from the adjustment of the etalon. It is noticed that the seven layer results are in general at higher resolution than the nine layer results. This is thought to be due in part to the greater knowledge of the importance of adjustment after the defects were observed as described in Chapter IV and also the adjusting mechanism made after the nine layer results.

Since the etalon function is of less importance at larger spacing it seems reasonable that the adjusting should not be as critical. However, since the fringes are broader this advantage may be nulified.

The adjusting is important and it would be valuable if the recorded function could be observed while adjusting since what

appears best to the eye when observing the plates may not give the desired instrumental profile. For example, it may be possible to adjust for maximum area in adjustment but yet come out with an asymmetrical instrumental profile. This would be easy if a fast scan and oscilloscope could be employed.

However, even with no improvements to the instrument it is at least as good as was hoped for when it was designed, and should be a useful instrument for many other studies.

VI CONCLUSION

A Fabry-Perot spectrometer was built and appears to be very satisfactory for auroral observations. It has enough sensitivity to be able to use the resolution required to obtain Doppler temperatures from the auroral green line. A much more extensive study should be made since actual temperatures have been obtained.

The temperatures obtained range from 222 ^{o}K to 720 ^{o}K, a much larger range than previously obtained with this method and in general much higher than the rotational temperature of the 3914A N_2 band. It would be valuable to try and correlate these temperatures with simultaneous 3914A N_2^+ temperatures.

The Fabry-Perot spectrometer lends itself well to a simultaneous study of the oxygen 5577A and 6300A lines, where a comparison of temperatures could be obtained. The instrument can be converted from one line to the other simply by changing an interference filter. At lower resolution the instrument, because of its high light gathering power, should be useful for observing the less intense emissions.

BIBLIOGRAPHY

Armstrong, E.B.	1956	'Airglow and Aurorae' ed. E.B. Armstrong and A. Dalgarno Pergamon Press, London.
Armstrong, E.B.	1959	J. Atmos. Terr. Phys. 13, 205
Babcock, H.D.	1923	Astro. Phys. J. 57, 209
Bernard, R.	1938	Zs. f. Phys. 110, 291
Cabannes, J., and Dufay, J.	1955	Comptes Rendus 240, 37
Cabannes, J., and Dufay, J.	1956a	Revue d'Optique 35, 103
Cabannes, J., and Dufay, J.	1956b	'Airglow and Aurorae' ed. E.B. Armstrong and A. Dalgarno Pergamon Press, London.
Chabbal, R.	1953	J. Rech. C.N.R.S. Number 29.
Chabbal, R.	1958	J. Phys. Rad. 19, 295
Dufay, J., Cabannes, J., and Gauzit, J.	1942	L'Astronomie 56, 149
Glasstone, S.	1940	Textbook of Physical Chemistry D. Van Nostrand Co., New York.
Hodder, B.H.	1958	J. of Sc. Ins. 35, 182
Hunten, D.M.	1953	Can. J. Phys. 31, 681-90
Jacquinot, P., and Dufour, C.	1949	J. Rech. C.N.R.S. Number 6, p. 91.
Jacquinot, P.	1954	J. Opt. Soc. Am. 44, 761
Jenkins, F.A., and White, H.E.	1950	'Fundamentals of Physical Optics' McGraw Hill, New York.
Jmaeff, A.	1959	B. Sc. Thesis, U. of S.

Karandikar, R.V.	1956	'Airglow and Aurorae' ed. E.B. Armstrong and A. Dalgarno Pergamon Press, London.
Karandikar, R.V.	1959	'Investigation of Interferometric Methods for Auroral Airglow Studies' Contract No AF19(604)-2172 Boston U. Phys. Res. Lab.
Meissner, K.W.	1941	J. Opt. Soc. Am. 31, 405
Rank, H.D.	1954	J. Opt. Soc. Am. 44, 5
Shepherd, G.G.	1953	M. Sc. Thesis, U. of S.
Shepherd, G.G., and Hunten, D.M.	1955	J. Atmos. Terr. Phys. 6, 328-35
Vegard, L.	1937	Phil. Mag. 24, 588
Wark, D.Q., and Stone, J.M.	1955	Nature 175, 254
Wark, D.Q.	1960	Astrophys. J. 131, 491
Wright, W.R. and Curtis, H.D.	1931	J. Opt. Soc. Am. 21, 159