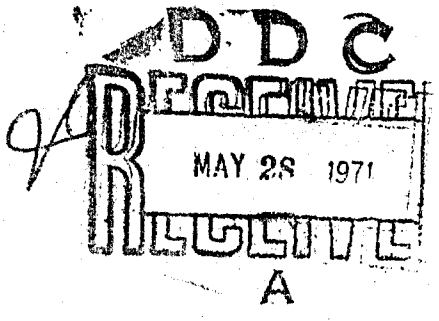
AFCRL-71-0019
5 JANUARY 1971
SPECIAL REPORTS, NO. 114

AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
L. G. HANSCOM FIELD, BEDFORD, MASSACHUSETTS

Aspen International Conference on Fourier Spectroscopy, 1970

GEORGE A. VANASSE
A.T. STAIR, JR.
DORAN J. BAKER

Editors



AIR FORCE SYSTEMS COMMAND
United States Air Force

AFCRL-71-0019
5 JANUARY 1971
SPECIAL REPORTS, NO. 114

OPTICAL PHYSICS LABORATORY PROJECT 8658, 7670

AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
L. G. HANSCOM FIELD, BEDFORD, MASSACHUSETTS

Aspen International Conference on Fourier Spectroscopy, 1970

GEORGE A. VANASSE
A.T. STAIR, JR.
DORAN J. BAKER*

Editors

* Utah State University, Electro-Dynamics Laboratories, Logan, Utah

AIR FORCE SYSTEMS COMMAND
United States Air Force

Foreword

It is a rare scientific meeting that brings together at one time and location the world's leading practitioners in a particular field. During the week of March 16–20, 1970, some 230 scientists attended the International Conference on Fourier Spectroscopy at the Aspen Institute in Colorado, the first truly international scientific meeting devoted exclusively to Fourier spectroscopy. Among those present were the pioneers in the field, those who have contributed most to the advance of the art, and those who are applying the art to a range of scientific endeavors.

The Air Force Cambridge Research Laboratories were proud to sponsor this Conference. AFCRL sponsors many scientific meetings each year, but our role in the International Conference on Fourier Spectroscopy was an especially rewarding one. Attending from AFCRL were 22 scientists, one group of whom reported on the Fourier Spectroscopic analysis of data from an event that occurred only a few days before the Conference—The 7 March 1970 solar eclipse. Other AFCRL scientists reported on the application of Fourier spectroscopy to the measurement of emission from remote infrared sources and of the transmission properties of the atmosphere.

I single out these AFCRL papers to point up the fact that for laboratories engaged in environmental research, such as AFCRL, this new spectroscopic tool is a powerful one indeed, bringing orders of magnitude improvement in the accuracy and resolution of sensor data. Even if we at AFCRL were simply users of Fourier spectroscopy, we would consider the sponsorship of such a Conference eminently worthwhile. But it is in those contributions that our own scientists have made to advance the art itself that AFCRL takes real pride.

The format chosen for the Conference is one that lends itself particularly well to meetings on highly specialized fields. The Conference was divided chronologically into two major parts. Tutorial sessions were held on Monday and Tuesday, followed by 3 days of invited and contributed research papers. A morning technical session, lasting until noon, was followed by a long afternoon break, with the participants reassembling at about 4:30 for 2 more hours of technical sessions. After dinner, informal discussions lasted until late at night.

This report on the Conference, of course, cannot convey the spontaneous and unrecorded insights that flavored the post-prandial sessions, nor the information informally exchanged, but the report does serve to give those who were not able to attend the Conference a complete state-of-the-art survey of Fourier spectroscopy.

The rapid advances made in Fourier spectroscopy—only 20 years to reach maturity—are recognized as among the most important advances yet made in the field of infrared spectroscopy.

DALE J. FLINDERS, Colonel, USAF
Commander
Air Force Cambridge Research Laboratories

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (*Corporate author*)

Air Force Cambridge Research Laboratories (OP)
L.G. Hanscom Field
Bedford, Massachusetts 01730

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP

3. REPORT TITLE

ASPEN INTERNATIONAL CONFERENCE ON FOURIER SPECTROSCOPY, 1970

4. DESCRIPTIVE NOTES (*Type of report and inclusive dates*)

Scientific. Interim.

5. AUTHOR(S) (*First name, middle initial, last name*)

George A. Vanasse
A. T. Stair, Jr.
Doran J. Baker Editors

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
5 January 1971	478	564

8a. CONTRACT OR GRANT NO.

b. PROJECT, TASK, WORK UNIT NOS. 8658-04-01 7670-05-01

c. DOD ELEMENT 61102F 62101F

d. DOD SUBELEMENT 681310 681000

9a. ORIGINATOR'S REPORT NUMBER(S)
AFCRL-71-0019

9b. OTHER REPORT NO(S) (*Any other numbers that may be assigned this report*)
Special Reports, No. 114

10. DISTRIBUTION STATEMENT

1. This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES
TECH, OTHER

12. SPONSORING MILITARY ACTIVITY
Air Force Cambridge Research Laboratories (OP)
L.G. Hanscom Field
Bedford, Massachusetts 01730

13. ABSTRACT

This report constitutes the proceedings of the Aspen International Conference on Fourier Spectroscopy. The tutorial lectures, which are also included in the proceedings, were intended to bring the participants to the point where the invited and contributed papers would be beneficial to all. These lectures began with a mathematical introduction and with a comprehensive outline of the technique of Fourier spectroscopy. The following specific topics were treated in more detail: signal-to-noise considerations, interferometers for Fourier spectroscopy, double-beaming techniques, refractometry, data handling and processing, and finally a speculative digression on the impact new developments in sources and detectors might have on Fourier spectroscopy.

The invited and contributed papers ran the gamut of high and low resolving power instrumentation and results; from resolutions of tens of wavenumbers to thousandths of wavenumbers. The topics covered included new instrumentation, new data handling and analysis techniques, advantages of Fourier techniques as demonstrated by recent results, theoretical considerations on general problems associated with the technique, and applications of the technique under adverse conditions.

Finally, it was decided to accept papers on multiplex techniques other than Fourier and/or interferometric. These papers were presented in the last session of the meeting and are also included in this volume.

DD FORM 1 NOV 65 1473

Unclassified

Security Classification

KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Fourier spectroscopy Interferometers Double-Beaming techniques Refractometry						

Abstract

This report constitutes the proceedings of the Aspen International Conference on Fourier Spectroscopy. The tutorial lectures, which are also included in the proceedings, were intended to bring the participants to the point where the invited and contributed papers would be beneficial to all. These lectures began with a mathematical introduction and with a comprehensive outline of the techniques of Fourier spectroscopy. The following specific topics were then treated in more detail: signal-to-noise considerations, interferometers for Fourier spectroscopy, double-beaming techniques, refractometry, data handling and processing, and finally a speculative digression on the impact that new developments might have on Fourier spectroscopy.

The invited and contributed papers ran the gamut of high and low resolving power instrumentation and results, and from resolutions of tens of wavenumbers to thousandths of wavenumbers. The topics covered included new instrumentation, new data handling and analysis techniques, advantages of Fourier techniques as demonstrated by recent results, theoretical considerations on general problems associated with the techniques, and applications of the techniques under adverse conditions.

Finally, it was decided to accept papers on recent multiplex techniques other than Fourier and/or interferometric. These papers were presented in the last session of the meeting and are also included in this report.

Contents

Contributed Papers (contd)

Contributed Papers (contd)

Exhibits

Opening Address

John N. Howard
Chief Scientist, AFCRL

On behalf of the Air Force Cambridge Research Laboratories, it is a pleasure to welcome the distinguished participants to this Conference on Fourier Spectroscopy and to wish you a successful meeting.

Speaking personally, I have three separate interests in this topic. First of all I have watched this scientific field grow in the last 20 years from an impractical daydream to its present state of spectacular success. Just after the war, I went to Ohio State to become educated in infrared physics. Our experimental apparatus was for the most part completely conventional. We did not have any fancy computers or electronics, but we did have newer bigger gratings and some of the improved infrared detectors that had been developed during the war. All around me students were writing theses on improved high resolution studies of various molecules. So I wrote a thesis using a simple Perkin-Elmer prism spectrometer and at the lowest resolution possible: the slits were wide open—2-mm wide. I was at Columbus when Marcel Golay gave his first paper on multiplexing with his multislit, and I heard Peter Fellgett—who was then at Michigan—first describe the multiplex gain. Brilliant ideas were flashing all around in those days, but I was incombustible. However, after leaving Ohio State I came to AFCRL where we were able to encourage infrared interferometry by sponsoring the research of many of the groups who are represented here at this meeting.

Secondly, I have been editor of *Applied Optics* since its inception in 1960, and that journal owes much of its prosperity to the renaissance in modern optics brought about not only by the laser but also by the availability of computers. Faint Raman lines that used to take days of exposure time with a mercury arc source are now recorded in milliseconds by using a laser, and magnificent spectra of molecular absorption in the atmospheres of faint stars and planets are now obtained by Fourier spectroscopy. Wavefront reconstruction, holography, and Fourier spectroscopy are some of the fields that *Applied Optics* has tried to provide a home for, and I have enjoyed presiding over some of the happy squabbles in these fields.

I have a third interest in the topic of this symposium because as most of you know the Research Library at the Air Force Cambridge Research Laboratories contains most of the scientific papers, correspondence, and other manuscript materials of the third Lord Rayleigh. Rayleigh was a friend of Michelson; they visited each other several times and also corresponded. Rayleigh gave Michelson much advice on how to interpret the interferograms Michelson was obtaining with his interferometer. But Rayleigh took no personal credit for this advice. He said the theory has been known for a long time: Fourier worked it out nearly one hundred years ago.

Aspen International Conference on Fourier Spectroscopy, 1970

Introduction

The Aspen International Conference on Fourier Spectroscopy (Aspen Conference) held March 16-20, 1970, in Aspen, Colorado, was the fourth conference wholly or partially devoted to the technique of Fourier spectroscopy. The Bellevue Conference in 1957 and the Orsay Conference in 1966 were sponsored by the Laboratoire Aimé Cotton of the Centre National de la Recherche Scientifique and the MISFITS Conference in 1964 was sponsored by the Mellon Institute in Pittsburgh, Pa. The Aspen conference was sponsored by the Air Force Cambridge Research Laboratories, but the papers presented at the Conference and presented herein do not necessarily reflect the views of either the editors or the sponsor.

The proceedings are made up of the text material from the tutorial sessions, which were given during the first 2 days of the conference, plus invited and contributed papers. The emphasis was placed on current applications, techniques, and data obtained rather than on subjects of purely spectroscopic interest. Various organizations had exhibits on display on March 18th; these exhibits are also described in this volume.

This conference was arranged and conducted through a contract with the Electro-Dynamics Laboratories of Utah State University. The organizing committee consisted of Dr. George A. Vanasse, AFCRL, Co-Chairman; Dr. A. T. Stair, AFCRL, Co-Chairman; Dr. Doran J. Baker, Utah State University, Secretary; Dr. Glen Smerage, Utah State University, Manager; Dr. Ely Bell, Ohio State University; and Dr. Larry Mertz, Smithsonian Astrophysical Observatory. Mr. King Woodward of the Institute for Humanistic Studies of Aspen was host for the Conference. Professor Pierre Jacquinot of C. N. R. S. gave the feature talk at the conference banquet applying the advantages of multiplexing to the administration of scientific research.

Tutorial Lectures

Contents

1. Fourier Spectroscopy: An Introduction

Ernest V. Loewenstein
Air Force Cambridge Research Laboratories,
Bedford, Massachusetts

Abstract

In this paper we start with an elementary discussion of the Michelson interferometer, leading into an extensive discussion of the mathematics involved in Fourier spectroscopy. Emphasis is placed on the use of the convolution theorem. The maximum allowable solid angle of light beam illuminating an interferometer is derived and compared with grating instruments. Special aspects of Fourier spectroscopy such as apodization, noise, and mathematical filtering are discussed.

1–1 THE MULTIPLEX PRINCIPLE

The superiority of Fourier spectroscopy over grating spectroscopy for high resolution work under low light level conditions has been incontrovertibly demonstrated by the Conneses in their near infrared planetary spectra[1]. There are many reasons why the Fourier method is inherently superior. The two most often quoted are the multiplex advantage (Fellgett) and the aperture advantage (Jacquinot). In addition, absolute wavenumber accuracy is guaranteed by the known wavelength used for carriage control, the physical apparatus is inherently simple, and both stray light and overlapping spectral orders are eliminated. The multiplex gain is the salient feature of Fourier spectroscopy, and we will commence with an elementary derivation.

Let us assume that the spectrum to be investigated extends from σ_1 to σ_2 wavenumbers, that the desired resolution is $\delta\sigma$, that the system is detector-noise limited, and the time available for the study is T. We define the number of spectral elements:

$$m = \frac{\sigma_2 - \sigma_1}{\delta\sigma}. \tag{1-1}$$

If we observe each element sequentially (as with a grating spectrometer) for a time T/m, the signal to noise ratio will be proportional to $(T/m)^{1/2}$, while if we observe each element for the entire time, T, the signal to noise for each element will be proportional to $T^{1/2}$. There is thus a gain of a factor of $m^{1/2}$ when all the spectral elements are observed concurrently, which is the multiplex gain. It sometimes happens, as in emission spectroscopy, that parts of the spectral range σ_1 to σ_2 contribute no energy to the signal. The effective number of spectral elements is then

$k<m$, but there remains a multiplex advantage as long as $k>1$.

There must, of course, be a method of coding the spectral elements so that they can subsequently be separated unambiguously. This method is provided by the two-beam interferometer that changes each wavenumber, σ, in the spectrum into an electrical frequency, f, according to the equation

$$f=v\sigma \tag{1-2}$$

where v is the rate of change of path difference. The superposition of all these frequencies is the interferogram, which is then reduced to a spectrum by means of a Fourier transformation. The Michelson interferometer (or one of its variants) is almost universally used for Fourier spectroscopy, and all discussions in this paper refer directly to it.

1-2 THE MICHELSON INTERFEROMETER I: ELEMENTARY CONSIDERATIONS

An elementary discussion of the Michelson interferometer serves as a convenient starting point to introduce many of the ideas we will need in Fourier spectroscopy and to lead us into the background mathematics that will be developed in Section 1-3. The interferometer and collimating optics are illustrated in Figure 1-1. In this initial discussion we will

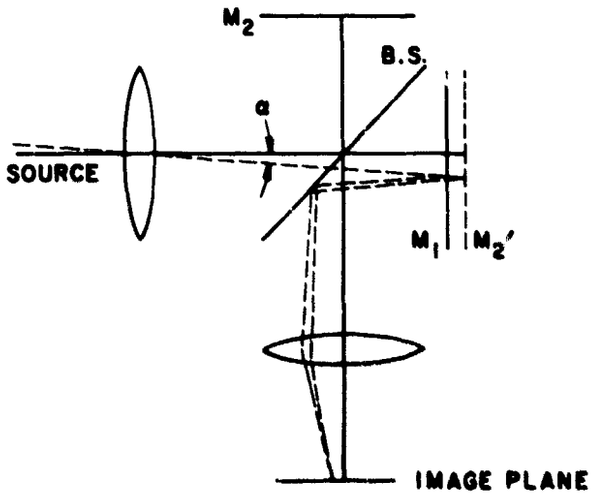


Figure 1-1. Michelson Interferometer. M_1 and M_2 are the end mirrors, M_2' is the image of M_2 as seen through the beamsplitter (BS)

assume an on-axis, quasi-monochromatic point source and a beamsplitter of negligible thickness, with (complex) amplitude transmittance and reflectance t and r, respectively. Let the incident wave be $A\exp[i(\omega t-2\pi x\sigma)]$. Then the net amplitude emerging from the interferometer in the direction of the detector is

$$A'_{\text{det}}=A(rt)[e^{i(\omega t-2\pi x_1\sigma)}+e^{i(\omega t-2\pi x_2\sigma)}] \tag{1-3}$$

where x_1 and x_2 are the round-trip distances from the beamsplitter to M_1 and M_2, respectively. The energy reaching the detector is:

$$E_{\text{det}}=|A'_{\text{det}}|^2=2A^2|rt|^2[1+\cos 2\pi(x_1-x_2)\sigma]. \tag{1-4}$$

Let $A^2=B(\sigma)\,d\sigma$; $|rt|^2=\varepsilon$, the beamsplitter efficiency; $x_1-x_2=x$, the path difference. Then

$$E_{\text{det}}=2\varepsilon B(\sigma)[1+\cos 2\pi\sigma x]\,d\sigma. \tag{1-5}$$

The interferogram is defined as the varying part of Eq. (1-5); i.e.,

$$dI(x)=2\varepsilon B(\sigma)(\cos 2\pi\sigma x)\,d\sigma \tag{1-6}$$

and we see immediately that the interferogram produced by a quasi-monochromatic line is a cosine function. A broad spectral range, then, requires an integral over σ:

$$I(x)=\int_0^\infty dI(x)=2\varepsilon\int_0^\infty B(\sigma)(\cos 2\pi\sigma x)\,d\sigma \tag{1-7}$$

which is the cosine Fourier integral of the spectrum. The recovery of the spectrum is then achieved by taking the inverse Fourier transform.

It is worthwhile to make a slight digression at this point to show, in an order-of-magnitude calculation, the relationship between the resolution, σ, and the maximum path difference, L, attained in the inter-

ferogram. We take as an arbitrary criterion that the minimum resolvable wavenumber interval occurs when there is a difference of one cycle of interference between two closely spaced lines. Thus, we have for one of the lines

$$L = m/\sigma \tag{1-8}$$

and for the other

$$L = \frac{m+1}{\sigma+\delta\sigma}. \tag{1-9}$$

Eliminating L between Eqs. (1-8) and (1-9), we have

$$\delta\sigma = \sigma/m, \tag{1-10}$$

and substituting from Eq. (1-8) we find that, for an order of magnitude criterion,

$$\delta\sigma \approx 1/L. \tag{1-11}$$

The resolution of the interferometer is inversely proportional to the path difference between the interfering beams. This is identical with the situation we find in the use of a diffraction grating, which gives its highest resolution when used at grazing incidence, where the path difference between the extreme interfering rays is a maximum.

Returning now to the question of reducing the interferogram, we see that we must discuss not only the Fourier transform but also sampling theory, because we must sample the interferogram to read it into the digital computer. The question of analog computation, which has received some attention in the past, need no longer concern us. The analog computer has severely limited accuracy and dynamic range compared to even the smallest modern digital computer. The advent of the fast Fourier transform has eliminated considerations of cost and computing time. The digital computer, furthermore, may be programmed to do much more than merely compute the Fourier transform. It may, for example, be used to compute line positions, or intensities, absorption coefficients, energy levels, or any of a host of other spectroscopic quantities. It may even be used to control the entire process from taking the interferogram to producing a graphical display of the results.

1-3 MATHEMATICS FOR FOURIER SPECTROSCOPY

1-3.1 The Fourier Integral

The Fourier integral may be defined by the pair of equations

$$f(x) = \int_{-\infty}^{\infty} F(\sigma) e^{i2\pi\sigma x}\, d\sigma \tag{1-12a}$$

$$F(\sigma) = \int_{-\infty}^{\infty} f(x) e^{-i2\pi x\sigma}\, dx \tag{1-12b}$$

or by the representation equation

$$f(x) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x) e^{-i2\pi x\sigma}\, dx \right] e^{i2\pi x\sigma}\, d\sigma. \tag{1-13}$$

The reciprocal quantities x and σ have dimensions of length and inverse length respectively. Fourier transform pairs will, with one exception, be denoted by using lower and upper case of the same letter; i.e., $F(\sigma) = F.T.\ [f(x)]$. The exception is that the interferogram and spectrum will be denoted by $I(x)$ and $B(\sigma)$, respectively. The meaning of Eq. (1-13) is that $f(x)$ may be represented by the process described, i.e., a "round trip" through the Fourier transform. The existence conditions are: (1) $f(x)$ must be absolutely integrable, i.e.,

$$\int_{-\infty}^{\infty} |f(x)|\, dx < M$$

where M is some finite number, and (2) $f(x)$ may have at most a finite number of finite discontinuities. At a point of discontinuity, it can be shown that the integral of Eq. (1-13) converges to $\frac{1}{2}[f(x+)+f(x-)]$, i.e., the midpoint of the jump. Certain obvious functions that do not have Fourier transforms are:

(1) A constant

(2) Any periodic function

(3) $f(x) = 1/x$.

A pictorial table of some elementary and useful Fourier transforms is given in Figure 1-2, and we use

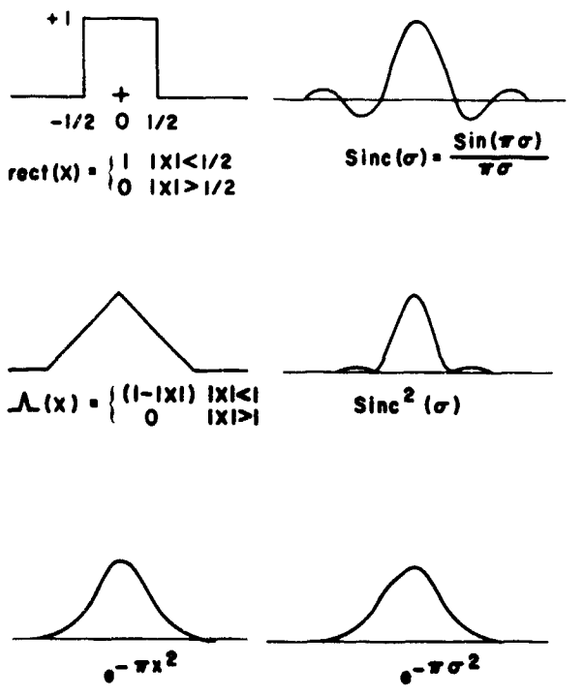


Figure 1-2. Some Useful Fourier Transform Pairs

this occasion to define these useful functions:

$$\text{rect}\,(x) = \begin{cases} 1 & |x| < \frac{1}{2} \\ 0 & |x| > \frac{1}{2} \end{cases} \tag{1-14}$$

$$\Lambda(x) = \begin{cases} 1-|x| & |x| < 1 \\ 0 & |x| > 1 \end{cases} \tag{1-15}$$

$$\text{sinc}\,(x) = \frac{\sin \pi x}{\pi x}. \tag{1-16}$$

The value of rect (x) at $x = \pm\frac{1}{2}$ is not defined, but we will only be using it in the form of Fourier transforms and thus will automatically get

$$\text{rect}\,(-\tfrac{1}{2}) = \text{rect}\,(\tfrac{1}{2}) = \tfrac{1}{2}.$$

The functions illustrated in Figure 1-2 are also an example of the fact that at least one of a Fourier pair has infinite extent. In the case of Fourier spectroscopy, a spectrum of finite extent produces an interferogram of infinite extent, which must of course be truncated, and the consequences of this truncation will be discussed presently.

There are three important theorems regarding Fourier transforms that we shall need upon occasion, and they are listed here without proof:

Shift theorem

$$FT[f(x+a)] = e^{i2\pi\sigma a}F(\sigma) \tag{1-17}$$

Scale change

$$FT[f(ax)] = \frac{1}{|a|}F\left(\frac{\sigma}{a}\right) \tag{1-18}$$

Rayleigh's Theorem

$$\int_{-\infty}^{\infty} |f(x)|^2\,dx = \int_{-\infty}^{\infty} |F(\sigma)|^2\,d\sigma. \tag{1-19}$$

(Rayleigh's theorem is the analog of Parseval's theorem for Fourier series). The proof of the first two is accomplished by a simple change of variable; the third is slightly longer and may be found in Bracewell[2].

1-3.2 Even and Odd Functions

An even function is one for which

$$E(x) = E(-x)$$

while an odd function has

$$O(x) = -O(-x).$$

Any complex function may be written as

$$f(x) = E'(x) + O'(x) + i[E''(x) + O''(x)]. \tag{1-20}$$

The Fourier transform of any even function reduces to a cosine transform, for an odd function it becomes a sine transform. Thus

$$F(\sigma)=FT[f(x)]=2\int_0^\infty E'(x)\cos 2\pi\sigma x\,dx$$
$$+2i\int_0^\infty E''(x)\cos 2\pi\sigma x\,dx$$
$$+2i\int_0^\infty O'(x)\sin 2\pi\sigma x\,dx$$
$$-2\int_0^\infty O''\sin 2\pi\sigma x\,dx.$$

If we now consider only real $f(x)$, such as the interferograms encountered in Fourier spectroscopy, we have

$$E''(x)\equiv O''(x)\equiv 0$$

and

$$F(\sigma)=2\int_0^\infty E'(x)\cos 2\pi\sigma x\,dx$$
$$+2i\int_0^\infty O'(x)\sin 2\pi\sigma x\,dx. \qquad (1\text{-}21)$$

The spectrum, thus, is Hermitian, or complex symmetric; i.e.,

$$F(\sigma)=F^*(-\sigma).$$

That is, no matter how badly distorted the interferogram may be, the spectrum derived from it is not worse that Eq. (1-21) (which is bad enough).

1-3.3 The δ Function

The δ function is best defined in terms of the important sifting property

$$\int_{-\infty}^{\infty}\delta(x)f(x)\,dx=f(0), \qquad (1\text{-}22)$$

provided that $f(x)$ is continuous at $x=0$. The normalization condition is added:

$$\int_{-\infty}^{\infty}\delta(x)\,dx=1. \qquad (1\text{-}23)$$

This function is often referred to as the Dirac δ-function, but it was not entirely an original idea with Dirac; physicists had long felt the need of a function that is large in a localized region and small everywhere else, to pick out the value of a field variable at one point. Dirac originally defined $\delta(x)$ as if it were a proper function with a value $f(x)$ assigned to every x. This procedure quickly encounters trouble with the mathematical formalities, but this can be largely avoided by using Eq. (1-22) to define the one property of $\delta(x)$ that we really need. (A discussion of the mathematical niceties is given by Papoulis[3].)

Two important properties of $\delta(x)$ easily shown by change of variable are:

$$\int_{-\infty}^{\infty}\delta(x-a)f(x)\,dx=f(a) \qquad (1\text{-}24)$$

and

$$\int_{-\infty}^{\infty}\delta(ax)f(x)\,dx=\frac{1}{|a|}\int_{-\infty}^{\infty}\delta(x)f\left(\frac{x}{a}\right)dx. \qquad (1\text{-}25)$$

Using the form of Eq. (1-24), we consider the Fourier transform

$$FT[\delta(x-a)]=\int_{-\infty}^{\infty}\delta(x-a)e^{i2\pi\sigma x}\,dx$$
$$=e^{i2\pi\sigma a}, \qquad (1\text{-}26)$$

which is a monochromatic, complex harmonic function. To get a real harmonic function we may use either of the following:

$$FT[\tfrac{1}{2}(\delta(x-a)+\delta(x+a))]=\cos 2\pi\sigma a \qquad (1\text{-}27a)$$

or

$$FT[-i/2(\delta(x-a)-\delta(x+a))]=\sin 2\pi\sigma a. \qquad (1\text{-}27b)$$

The inverse Fourier transform, however, does not exist, since the results in Eqs. (1-26) and (1-27) are not absolutely integrable. The δ-function, therefore, is not strictly suited to Fourier theory, unless we somehow terminate its harmonic transform. We shall consider the consequences of doing this, after discussing the convolution.

1-3.4 Convolution and Autocorrelation

We define the convolution of two functions

$$h(x) = f(x) * g(x) = \int_{-\infty}^{\infty} f(u)g(x-u)\,du \quad (1\text{-}28)$$

and the autocorrelation of a function

$$f(x) \bigstar f(x) = \int_{-\infty}^{\infty} f(x')f(x'+x)\,dx'. \quad (1\text{-}29)$$

We shall consistently use the asterisk and the five-pointed star to represent convolution and autocorrelation, respectively. The autocorrelation for even functions is easily shown to be self-convolution of the function. These two processes are important to us because there is a useful theorem regarding the Fourier transform of each. The convolution theorem states

$$FT[f(x) * g(x)] = FT[f(x)] \cdot FT[g(x)], \quad (1\text{-}30)$$

i.e., the Fourier transform of a convolution of two functions is the product of the Fourier transforms of the individual functions. Multiplication and convolution may thus be interchanged, at the cost of performing some Fourier transforms. The theorem regarding the autocorrelation is the Wiener-Khinchine theorem, which states that the Fourier transform of the autocorrelation of a function is its power spectrum. This provides the necessary link between the interferogram and the spectrum: for an interferogram is the autocorrelation of the incident wave amplitude.

It is worth spending a little time to clarify graphically the meaning of the convolution, as contrasted with multiplication. In multiplication the product of two functions is obtained simply from the product of the ordinates in the region of overlap, when one is placed over the other with their origins coincident. Convolution involves a displacement of the origin of one of the functions with respect to the other and, in fact, it is the distance between the origins that is the independent variable of the convolution, $h(x)$. The entire procedure of convolution may be outlined as follows:

(1) Reverse one of the functions, say $f(x)$. (It is immaterial which one is reversed, since it is elementary to show that $f * g = g * f$.)

(2) Displace the origin of $f(x)$ to the left by some sufficient distance; call it x_0.

(3) Lay the reversed, displaced $f(x)$ over $g(x)$.

(4) Multiply the functions in the overlap region and integrate the product.

(5) This forms the convolution at $x = -x_0$.

(6) Move the displaced function to the right by a distance Δx, and repeat the procedure.

The "sufficient" distance, mentioned in Step 2, is obvious in the case of functions of finite extent in x; it is the minimum distance necessary to produce zero overlap. Figure 1-3 illustrates progressively the convolution of two rect functions. While it is true that this example is equally the autocorrelation function, its simplicity and clarity commend it to us.

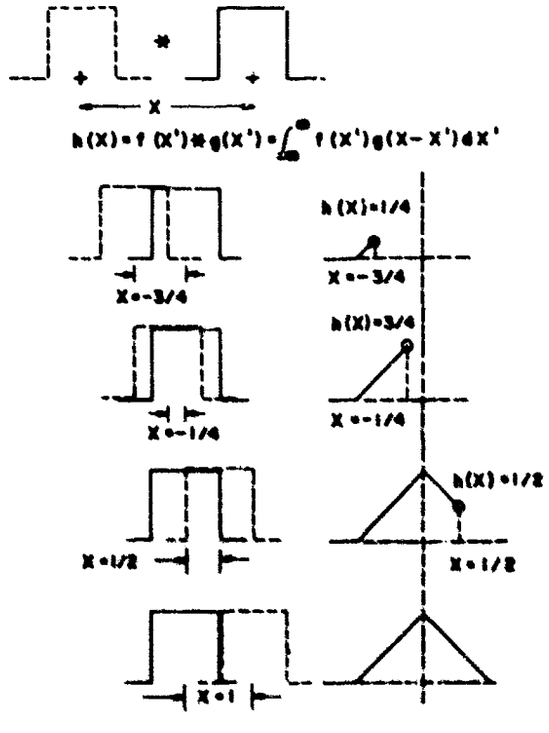


Figure 1-3. Illustration of the Convolution Process, Using the Convolution of Two Rect Functions. The progress of the convolution function, $h(x)$, is shown as the convolvant is slid along

The interested reader will find numerous other examples illustrated in Chapter 3 of Bracewell[2].

It is not always so easy to see through the convolution process, and we now turn to an example where the convolution theorem is useful. Let us inquire for the convolution.

$$h(x) = \text{sinc}\,(x) * \text{sinc}\,(x).$$

These are functions of infinite extent with many "wiggles", and graphical methods will obviously lead to confusion. We may take the Fourier transform

$$H(\sigma) = FT[\text{sinc}\,(x) * \text{sinc}\,(x)]$$

and, by the convolution theorem, Eq. (1-30),

$$H(\sigma) = FT[\text{sinc}\,(x)] \cdot FT[\text{sinc}\,(x)].$$

We already know that the Fourier transform of sinc (x) is rect (σ) (see Figure 1-2), so

$$H(\sigma) = \text{rect}\,(\sigma) \cdot \text{rect}\,(\sigma) = \text{rect}^2\,(\sigma) = \text{rect}\,(\sigma)$$

and

$$h(x) = FT[H(\sigma)] = \text{sinc}\,(x).$$

We have the rather surprising conclusion that sinc * sinc = sinc (and also incidentally that sinc (x) is its own autocorrelation). This is less surprising when we consider that the frequency content of the sinc function is given by the rect function, which is constant up to some cutoff. Multiplying by another rect of the same width does not alter the frequency content and thus gives back the original function in the x domain.

In the examples above we have illustrated the convolution of two functions of equal width. The more common situation is that one of the convolvants is considerably narrower than the other. Consider, for example, the convolution of a rect function with an arbitrary function that has features that are narrow compared to the width of the rect function. The effect of the convolution is to smooth out the narrow features. This is the "blurring" or "running average" property of the convolution. The average we get is weighted by the shape of the narrow convolvant, and considerable distortion can result if its shape is sufficiently outlandish.

As a final example of the convolution theorem we will illustrate the effect of truncating the interferogram. It will be recalled that the interferogram arising from a pair of δ functions centered at $\pm\sigma_0$ is $\cos 2\pi\sigma_0 x$. Now assume this is truncated by rect $\frac{x}{L}$. The interferogram then becomes

$$I'(x) = \cos 2\pi\sigma_0 x \text{ rect}\left(\frac{x}{L}\right)$$

and the Fourier transform is:

$$B'(\sigma) = \frac{L}{2}[\delta(\sigma-\sigma_0) + \delta(\sigma+\sigma_0)] * \text{sinc}\,(L\sigma).$$

The sinc function is referred to as the scanning function for rectangular truncation. Its convolution with the δ function yields sinc function spectral lines. These of course are considerably distorted from the lines we started with, since the δ function has zero width and no side lobes. The width of the sinc function at half maximum, which may be taken as a measure of the resolution of the interferometer, is $1.2/L$. This is in consonance with our previous order-of-magnitude calculation that yielded $1/L$. The side lobes of the sinc function are particularly undesirable, for when convolved with a spectral feature narrow compared to $1.2/L$ they can produce spurious oscillations, which are exemplified in the well-known Gibbs' phenomenon. These side lobes may be suppressed by the process of apodization that is discussed in Section 1.6 of this paper.

1-3.5 Sampling and Replicating, the Shah Function

The sampling function, known as $\text{ш}(x)$ (shah), or the δ function comb, is defined as

$$\text{ш}(x) = \sum_{n=-\infty}^{\infty} \delta(x-n) \qquad (1\text{-}31)$$

which is a series of δ functions at the integers. It is quite obvious that

$$\text{ш}(x+m) = \text{ш}(x) \qquad (m = \text{any integer})$$

and it is not difficult to show that

$$\text{ш}(ax) = \frac{1}{(a)}\,\text{ш}\left(x-\frac{n}{a}\right). \qquad (1\text{-}32)$$

Less obvious, but of paramount importance for our study of Fourier spectroscopy, is the fact that ш is its own Fourier transform:

$$FT[\text{ш}(ax)] = \frac{1}{a}\,\text{ш}(\sigma/a). \qquad (1\text{-}33)$$

This is shown as follows:

$$FT[\text{ш}(ax)] = \frac{1}{|a|}\sum_{n=-\infty}^{\infty}\int_{-\infty}^{\infty}\delta\left(x-\frac{n}{a}\right)e^{i2\pi x\sigma}\,dx$$

$$= \frac{1}{|a|}\sum_{n=-\infty}^{\infty} e^{i\left(2\pi\sigma\frac{n}{a}\right)}.$$

The right-hand side is familiar from the theory of diffraction gratings or Fabry-Perot interferometers as a series of spikes of frequency $1/a$; since the sum truly extends to infinity, the spikes in this case are infinitely sharp.

We shall be concerned with ш not only as a sampling but also a replicating function. These properties are illustrated in Figure 1-4. Sampling is

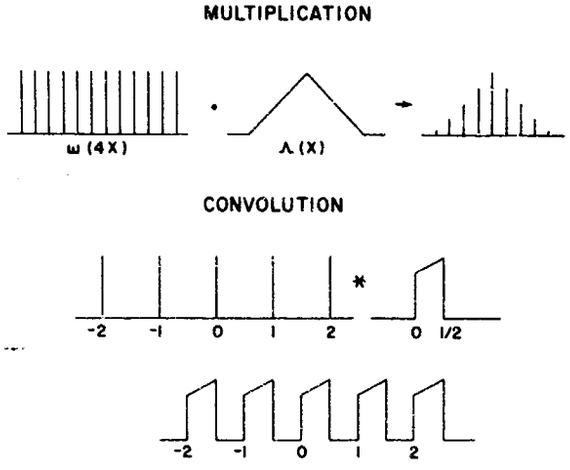


Figure 1-4. The Use of the ш Function for Sampling (by multiplication) and Replicating (by convolving). There is an understood integral over the infinite range in the multiplication process

accomplished by multiplication

$$\text{ш}(x)\,f(x) = \sum_{n=-\infty}^{\infty} f(n)\,\delta(x-n) \qquad (1\text{-}34)$$

and replication by convolution

$$\text{ш}(x) * f(x) = \sum_{n=-\infty}^{\infty} f(x-n). \qquad (1\text{-}35)$$

The difference between these two processes is the sliding property of the convolution, as has been discussed above.

When sampling an interferogram (or any function) it is of primary importance to know what sampling frequency is needed. This information is supplied by the sampling theorem that will be illustrated here, rather than rigorously derived. Assume we have a spectrum $B(\sigma)$ extending from 0 to σ_{max}, as illustrated in Figure 1-5. The interferogram is $I(x)$, and it is

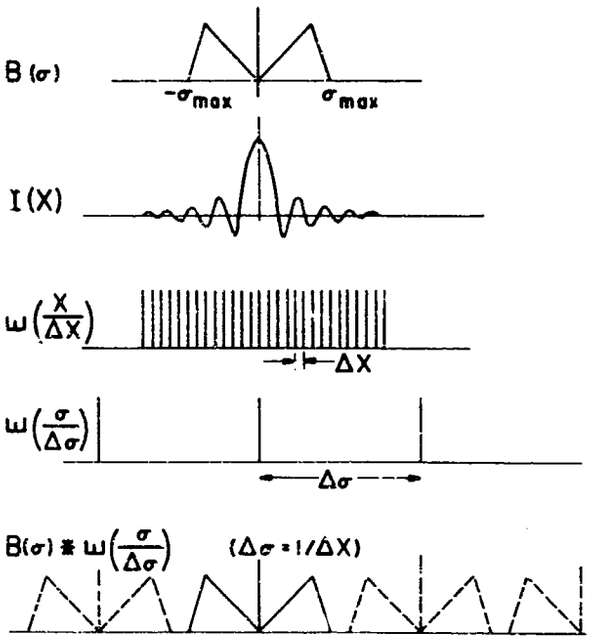


Figure 1-5. Illustrating the Replication of the Spectrum Arising From the Sampling of the Interferogram. The last line is used to show that the sampling interval in the interferogram must be $\Delta x = \frac{1}{2}\sigma_{max}$

sampled at intervals Δx; the sampled interferogram is

$$I'(x) = \text{ш}\left(\frac{x}{\Delta x}\right) \cdot i(x).$$

The spectrum derived from the Fourier transform of this interferogram is

$$B'(\sigma) = (\Delta x)\text{ш}\left(\frac{\sigma}{\Delta\sigma}\right) * B(\sigma)$$

$$(\Delta\sigma = 1/\Delta x).$$

It is clear from Figure 1-5 that in order to avoid overlap we must assure that $\Delta\sigma \geq 2\sigma$ max, which makes

$$\Delta x \leq \frac{1}{2\sigma_{max}}. \qquad (1\text{-}36)$$

This is the basic sampling theorem; i.e., that we must sample at a rate equal to the reciprocal of twice the highest frequency present in the record in order to avoid overlapping, or aliasing as it is called in communication theory. (This sampling frequency is called the Nyquist frequency in electrical engineering.)

In case the spectrum is band limited, i.e., extends from σ_1 to σ_2, where $\sigma_1 \neq 0$, we may be able to realize a saving in sampling, as illustrated in Figure 1-6.

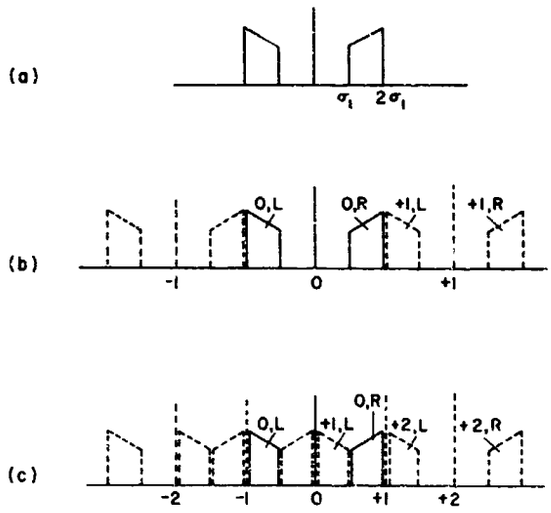


Figure 1-6. The Replicated Spectra for the Band Limited Sampling Theorem are Shown for the Case Where the Sampling Interval in (b) was $\Delta x = 1/\sigma_{max}$ and in (c) $\Delta x = 1/2(\sigma_2 - \sigma_1)$. Each vertical line is one element of the ш function and is given an ordinal number that is assigned starting from zero at the center of the original spectrum. There is a pair of replicated spectra about each of these elements, and each replicated spectrum is denoted by the ordinal number of the element it belongs to (0, 1, 2, . . .) and the letter L or R, depending upon whether it is the left or right member of the pair. The purpose of the illustration is to show that, in the case $\sigma_2 = 2\sigma_1$, halving the ordinary sampling interval produces no overlap.

There we have a spectrum extending from σ_1 to $2\sigma_1$ and we see that the space from 0 to σ_1 may be filled in with a replicated spectrum without incurring any overlap. This is an illustration of the band-limited version of the sampling theorem that provides that if the spectrum is limited to the band (σ_1, σ_2), the interferogram may be sampled at the rate

$$\Delta x = 1/(2(\sigma_2 - \sigma_1)), \qquad (1\text{-}37)$$

but with the important auxiliary condition that

$$\sigma_2 = h(\sigma_2 - \sigma_1) \text{ where } h = \text{an integer}, \qquad (1\text{-}38)$$

i.e., that the highest frequency be an integral multiple of the bandwidth. (A few minutes spent in making a drawing similar to Figure 1-6 in which Eq. (1-38) is not satisfied will serve to convince the reader that this is so.)

1-4 THE MICHELSON INTERFEROMETER, II: ETENDUE GAIN AND APERTURE EFFECTS

1-4.1 Size of the Aperture in a Michelson Interferometer

In Section 1-2 we discussed the Michelson interferometer illuminated by a parallel bundle of radiation emanating from a point source. In the real situation an extended source is used and we have rays that are not parallel to the axis. This produces the well-known circular fringe pattern when the planes of M_1 and M_2' are parallel (Figure 1-1). The diameter of the rings is a maximum at zero path difference, and we will concentrate our attention on the central spot to answer the question: what is the largest usable aperture the interferometer may subtend.

The ring pattern is illustrated in Figure 1-7 for the

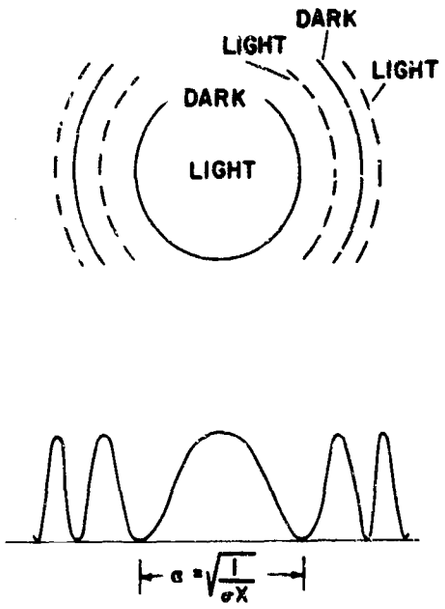


Figure 1-7. Upper Drawing: Ring Pattern of Michelson Interferometer Illuminated by Quasi-monochromatic Radiation With Finite Solid Angle; Lower Drawing: Intensity Variation Across the Pattern Illustrated Above

situation where the intensity is a maximum at the center (this illustration still applies to quasi-monochromatic illumination). As the path difference is increased, the rings shrink and the intensity at any point varies sinusoidally from a maximum to a minimum. We must therefore limit the aperture to include only one fringe, otherwise the interferogram will exhibit no variation of intensity with path difference.

The shrinking of the ring pattern with increasing path difference requires that we set the aperture to the size of the central spot at maximum path difference. The revision of Eq. (1-5) for the fringe pattern to include off axis rays is

$$dE = 2\mathcal{E}B(\sigma)\, d\sigma[1+\cos(2\pi\sigma x \cos\alpha)]\, d\Omega, \quad (1\text{-}39)$$

where α = the angle between the ray and the axis and $d\Omega$ is the solid angle element.

If we choose the path difference so that the center of the pattern has an intensity maximum, then the position of the first minimum may be computed from Eq. (1-39). Using the small angle approximation, we treat the argument of the cosine function

$$2\pi\sigma x \cos\alpha \cong 2\pi\sigma x\left(1-\frac{\alpha^2}{2}\right).$$

The phase difference between the central ray and the ray of the first intensity minimum is π:

$$2\pi\sigma x\left(\frac{\alpha_1^{\,2}}{2}\right) = \pi$$

$$\alpha_1^{\,2} = \frac{1}{\sigma x}.$$

The solid angle subtended from the center to the first intensity minimum for $x = L$ and $\sigma = \sigma_{max}$ is then

$$\Omega_1 = \pi\alpha_1^{\,2} = \frac{\pi}{L\sigma_{max}}. \quad (1\text{-}40)$$

Another effect on the interferogram arises from the integration of Eq. (1-39). Using the small angle approximation, and setting $\Omega = \pi\alpha^2$, the integration over solid angle yields:

$$E = 2\mathcal{E}B(\sigma)\, d\sigma\Omega \left[1+\operatorname{sinc}\left(\frac{\sigma x}{2\pi}\right)\cos\left(2\pi\sigma x_0\left(1-\frac{\Omega}{4\pi}\right)\right)\right]. \quad (1\text{-}41)$$

The interferogram function is defined as the second term on the right. There are now two distinct new features: the interferogram is modulated by a sinc function and the wavenumbers are shifted by an amount dependent upon the solid angle. The modulating sinc function has its first zero at

$$\Omega = \frac{2\pi}{\sigma x}.$$

If the interferometer is driven beyond the first zero of the modulating sinc function, the phase of the fringes is reversed and energy is removed from the spectrum rather than added to it as the path difference is increased. The absolute maximum aperture we may use is therefore:

$$\Omega_0 = 2\pi/L\sigma_{\max}. \tag{1-42}$$

This aperture is twice that of Eq. (1-41), but since the fringe contrast is attenuated severely toward the end of the interferogram, there is an effective apodization (to be discussed further below) that broadens the scanning function in the spectrum and the gain may be marginal. If Eq. (1-41) is used, the fringe contrast at the maximum path difference (for $\sigma_{\max}$) is 0.64 that at the center, which is not a severe apodization. If the larger solid angle is used, the fringe contrast (again for $\sigma_{\max}$) goes to zero, the scanning function is no longer sinc (σL) but a broader function, and it is wavenumber dependent, becoming narrower with decreasing wavenumber. For this reason, the aperture is conventionally set in accordance with Eq. (1-41).

1-4.2 Comparison of Etendue for Michelson and Grating Spectrometers

We define the étendue of an optical system as

$$E = A\Omega$$

where A = area of the collimator, and Ω the solid angle subtended by the detector.

The étendue is in general constant for an optical system (i.e., it cannot be increased), and it determines the amount of light that can be transmitted by the system. It is therefore sometimes referred to as "throughput" or "light-grasp" or even by the unhappy term "luminosity".

We have already determined the solid angle for the Michelson interferometer, which we will now rewrite using the following relations:

$$\delta\sigma = 1/2L$$

$$R = \sigma/\delta\sigma$$

($\delta\sigma$ = resolution, R = resolving power).

Substituting these into Eq. (1-42) yields

$$\Omega_M = \frac{2\pi}{R}. \tag{1-43}$$

To arrive at a corresponding equation for a grating spectrometer we note that the solid angle subtended by the exit slit is

$$\Omega_G = \frac{wl}{f^2}$$

where w and l are the width and length of the slit, respectively, and f is the collimator focal length. For a resolution $\delta\sigma$ and dispersion $d\theta/d\sigma$, the exit slit width is given by

$$w = f\frac{d\theta}{d\sigma}\,\delta\sigma = f\frac{d\theta}{d\sigma}\,\frac{\sigma}{R}$$

making

$$\Omega_G = \frac{l}{f}\,\frac{1}{R}\,\frac{d\theta}{d\sigma}\,\sigma.$$

It can easily be shown, using the grating equation, that

$$\frac{d\theta}{d\sigma}\sigma = \tan\theta \approx 1.$$

(We assume that θ is not grossly different from 45°—i.e., that it is not very close to 90° or 0°. If it were we would have either no energy or no dispersion.) This makes

$$\Omega_G = \frac{l}{f}\frac{1}{R} \tag{1-44}$$

and we compare the étendues

$$E_M = A\Omega_M = \frac{2\pi A}{R} \tag{1-45a}$$

$$E_G = A\Omega_G = \frac{l}{f}\frac{A}{R}. \tag{1-45b}$$

Even for a very fast grating spectrometer, l/f does not exceed $\frac{1}{30}$, which makes the étendue of the Michelson interferometer better than the grating by a factor of 200, for equal collimator area and resolving power, all other things being equal. This is the genesis of the so-called throughput gain of the Michelson interferometer, which is a direct result of the cylindrical symmetry of the interferometer. To realize an aperture gain from the use of the interferometer, it is important to meet the conditions of equal area and equal resolving power. It is also important that the detector be able to accept the added solid angle available from the interferometer. While this is usually the case, it requires attention in the design of the optical system.

1-5 SPECTRAL RECOVERY

Although the spectrum is obtained in principle by a Fourier transform of the interferogram, various factors intervene to make the recovered spectrum an imperfect representation of the true spectrum. The most important ones are:

1. Aperture effect
2. Tilt and aberrations
3. Truncation
4. Phase and compensation error
5. Noise.

Aperture effect has already been discussed above; it produces a shift of the computed wavenumbers and a reduction in contrast of the interferogram with increasing path difference. "Tilt" refers to failure of the movable mirror to translate strictly parallel to itself, and its effect is to change the contrast in the interferogram as some unpredictable function of the path difference. The subject of tilt compensation is taken up by Steel in Chapter 3. Aberrations in the optical system can produce asymmetric interferograms as a result of distortion of the interference fringe pattern (Figure 1-7). This destroys the cylindrical symmetry of the Michelson interferometer and consequently reduces the étendue. If the asymmetry is small, it can be corrected by the procedure outlined below.

1-5.1 Truncation; Phase and Compensation Error

Let us assume that a less-than-ideal interferometer produces an interferogram, $I(x)$, that, as a result of imperfect compensation for the beamsplitter, may not be an even function of x. The sampled interferogram that goes to the computer is then

$$I'(x) = I(x)T(x)\sqcup\!\sqcup(x+\epsilon). \tag{1-46}$$

The spectrum recovered by a cosine Fourier transform will be

$$B'(\sigma) = [B(\sigma) * t(\sigma)]e^{i\phi(\sigma)} \tag{1-47}$$

where $B(\sigma)$ is the true spectrum, $\phi(\sigma)$ is the phase function resulting from both the compensation error and the phase error ϵ in $\sqcup\!\sqcup(x+\epsilon)$. The latter arises when the sampling signal is not synchronized with the interferogram to produce a sample at exactly zero path difference. The effect of a non-zero phase function is to produce an asymmetric scanning function that not only distorts the observed spectral lines but also modifies the baseline of the spectrum, thus destroying the photometric accuracy of the measurements. $T(x)$ represents the truncating function that terminates the interferogram at some length, L. The simplest truncating function is rect (x/L), and we have already seen that this produces a scanning function sinc (Lx). This may be modified (but not eliminated) by multiplying the interferogram by another function, $A(x)$, called an apodizing function, as will be discussed in the next Section.

The phase function, on the other hand, may be eliminated in a broad range of situations. One possibility of course is to take the interferogram over positive and negative values of path difference, and compute both sine and cosine transforms. The amplitude of the spectrum is then truly free of phase distortions, provided the range of the original phase errors is small with respect to the total length of the interferogram. There are several reasons why this is not a desirable procedure. The time to take the interferogram, the total path difference required of the interferometer, and the total memory capacity required of the computer are all doubled. The computer time required for the fast Fourier transform is increased by a factor of a little more than 2. (If the interferogram has N points between zero and L, the double-sided transform requires a time $2+2/\log_2 N$ more than the single-sided one.) Finally, the process of squaring to calculate the amplitude changes the noise from a function randomly fluctuating about zero to one that fluctuates about a positive value, thus raising the effective noise level in the spectrum.

We therefore turn our attention to methods for determining the spectrum from a cosine Fourier transform alone. It is clear that if we had knowledge of $\phi(\sigma)$ we could correct Eq. (1-47) and obtain the spectrum

$$B''(\sigma) = [B(\sigma) * t(\sigma)] = B'(\sigma) \cdot e^{-i\phi(\sigma)} \qquad (1\text{-}48)$$

or, alternatively, we would correct the interferogram of Eq. (1-46) by the procedure

$$I''(x) = I'(x) * FTe^{-i\phi(\sigma)} \qquad (1\text{-}49)$$

and the FT of $I''(x)$ would yield $B''(\sigma)$. Either one of these methods works, although the latter is in more common use. The phase function $\phi(\sigma)$ is determined from the sine and cosine transforms of a short section of interferogram symmetric about zero path difference. If the phase function is a result of a failure to phase the sampling function properly, then it will be linear in wavenumber. If compensation error or aberrations are at fault, the phase function will in general not be linear. If the compensation error is very bad, then it may be difficult to produce a proper phase correction, because it will be difficult to find a point that can properly be labelled zero path difference.

1–5.2 Noise

There are four principal sources of noise in Fourier spectroscopy (and any other kind as well):

(1) Source
(2) Detector
(3) Scintillation
(4) Digitization.

Since there is a paper in this series by Sakai devoted to noise, we will limit this discussion to a few remarks about each kind of noise.

The multiplex gain of Fourier spectroscopy was derived in Section 1–1 on the assumption that the detector is the limiting source of noise for the system. If this is not so, then some other justification must be supplied for using Fourier spectroscopy (such as aperture gain). We will thus give no further attention to the first two noise sources.

Scintillation noise arises in the medium intervening between source and detector, and is most commonly observed when a long path through a gas or the atmosphere is involved, such as in astronomical work. Scintillation effects can be overcome to some extent by adjusting the interferogram frequency range to lie outside the range of scintillation frequencies, which are usually limited to some well defined band. The effects of scintillation noise may also be suppressed to a large extent by a scheme known as internal modulation. Instead of chopping the signal with a rotating blade or some similar means, chopping is accomplished by oscillating the path difference by an amount equal to one-half wavelength of the central wavenumber of the optical band reaching the detector. This has the effect of chopping the cosine dependent term in Eq. (1-40) without modulating the constant term, thus eliminating any change of level in the interferogram. It has been used with great success by Connes.[1]

Digitization noise arises from two sources: (1) unequal path difference between successive samples, and (2) the effect of the minimum detectable increment (quantification) of the digital voltmeter. The accuracy of the interval between successive readings is exactly as important as the accuracy of ruling a diffraction grating. It is a well-known fact that a grating with random errors produces fog or noise, and one with periodic errors produces ghosts; so it is in Fourier spectroscopy. For the far infrared it may be possible to rely on a high quality micrometer screw to provide the position readout that triggers the data acquisition system, but for any other region, either a good moiré system or an auxiliary measuring interferometer is indispensible.

The dynamic range of the interferogram is very large when a broad spectral range is being studied, i.e., the peak value at zero path difference is large compared to the oscillations in the remainder of the interferogram. The digitizing system must have adequate dynamic range to handle both the central maximum and the smaller oscillations without seriously compromising the signal-to-noise ratio. As

a rule of thumb we may apply the criterion that the dynamic range of the digitizing system must be at least equal to the signal-to-noise ratio in the interferogram at zero path difference.[4]

One final point on the subject of noise merits mention in this discussion, and that is with regard to the sampling theorem. The sampling theorem requires that we sample twice per highest frequency present in the record, regardless of whether that be a signal or noise frequency. Since a low-pass RC filter is used in many recording systems, it is usually true that the noise bandwidth is greater than the signal bandwidth. The filter time constant is generally chosen long enough to pass the highest signal frequency without appreciable attenuation or phase shift. The slow rolloff of an RC filter permits higher noise frequencies to come into the record than the signal frequencies present, and the sampling interval must be chosen accordingly.

1-6 SPECIAL TECHNIQUES OF FOURIER SPECTROSCOPY

There are three special aspects of Fourier spectroscopy to which we will give brief separate mention in this Section:

1. Apodization
2. Mathematical filtering
3. Refractometry

The first two are computational techniques, while the latter is essentially an experimental modification of the conventional spectroscopic setup.

Apodization[6] was mentioned earlier in the discussion of the scanning function. If the sinc scanning function is not to our liking it may be modified by the expedient of multiplying the right side of Eq. (1-46) by another function $A(x)$. The usual apodizing function, $A(x)$, is an even function that has value unity at $x=0$ and zero at $X=\pm L$. The new scanning function then is the Fourier transform of $A(x)$. The original truncating function, $T(x)$, has no influence on the new scanning function, since it does not alter the frequency content of $A(x)$ in any way. (See the remarks about the convolution of two sinc functions in Section 1-3.4.) The copious oscillations of the sinc function are a consequence of the square corners of the truncating function (rect (x)), which indicates that to reduce these oscillations we must choose for $A(x)$ a function that varies more smoothly. If we choose $A(x)=A(x)$, the scanning function becomes $\text{sinc}^2(x)$ (Figure 1-2), which has considerably smaller side lobes than sinc (x). It has, however, twice the width of the original scanning function for the same total interferogram length. The $(\text{sinc})^2$ function is also the slit function of a diffraction-limited grating spectrometer, and may have some advantages for comparison purposes. Whether it has merits in and of itself is a matter for individual workers to determine in their own situations.

A number of other apodizing functions have been discussed and used (see Ref. 2), but there is room for considerable discussion whether any apodization is necessary at all—at least under laboratory conditions. Apodization is basically a trade-off between the width of the scanning function (resolution) and its smoothness. The smoothness of the scanning function affects the ability to detect a weak line close to a strong one, but if the resolution sacrificed in reducing the oscillations of the scanning function is enough to blend the lines that were to be resolved, then nothing has been gained. Perhaps the best reason for applying apodization in laboratory situations is that it is simpler than exercising any other control over line widths (such as adjusting gas pressure). Its utility in other applications, such as astronomical Fourier spectroscopy, is more clear cut, as the conditions of excitation are not at the experimenter's disposal.

Mathematical filtering[7] is a technique for altering the frequency content of an interferogram prior to performing the Fourier transform. This permits a two-fold saving: a reduction in the number of samples required (saving on computer storage), and a reduction in computer time required to do the transformation. The method is based on the fact that reducing the spectral bandwidth may be accomplished by convolving the interferogram with a sinc function, which is equivalent to multiplying by a rect function in the spectrum. The processes in the x and σ domains are illustrated in Figure 1-8. The utility of the pro-

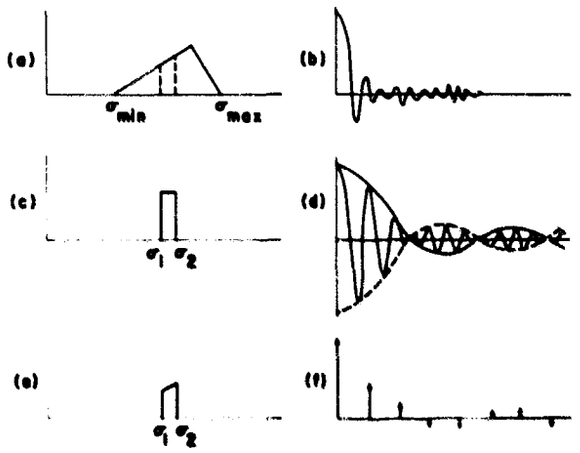


Figure 1-8. Mathematical Filtering: (a) the Original Spectrum and (b) the Original Interferogram. It is desired to limit the spectrum to the range (σ_1, σ_2), which may be accomplished by multiplying the spectrum by the rect function in (c) or convolving the interferogram with (d), the FT of the rect function. The FT of the rect function is a cosine "carrier" modulated by a sinc function envelope. The sinc function is characteristic of the width of the rect function and the "carrier" frequency depends upon the distance of the rect function from the origin (shift theorem). The resulting spectrum is shown in (e), and the sampled interferogram in (f). The sampling rate in the interferogram is determined by the sinc function frequency and not the "carrier" frequency, as a result of the band limited version of the sampling theorem

cedure lies in the fact that it may be necessary or desirable to record an interferogram with a broader range of signal frequencies than the experimenter intends ultimately to use. This could arise if appropriate filters were not available, or, in the case of a non-recurring event, when the best spectral region cannot be predicted.

The technique of refractometry[8,9] is derived from the familiar refractometers of the visible region, but with a difference. The sample is placed in one arm of the interferometer and an interferogram is taken, which will now assuredly be asymmetric. The phase curve contains the refractive index information and the absorption coefficient may be determined from the amplitude spectrum. The precision possible in this technique is a significant advance over other available techniques for determining far infrared optical properties. (Chapter 5 by Bell is devoted to this subject.)

1-7 COMPUTATION IN FOURIER SPECTROSCOPY

The overriding factor in computation for Fourier spectroscopy today is the fast Fourier transform, or Cooley-Tukey algorithm,[10] which has changed the computational problem from one of cost to one of finding a computer with sufficiently large memory capacity to do the desired transforms. The time required for the FFT is proportional to $N \log_2 N$ compared to N^2 for the conventional method (N = number of points transformed). An actual interferogram of 4096 points can be transformed in 14 sec by the FFT, compared to 2734 sec (45 min) by orthodox methods.[4] This fact has also served to make irrelevant any discussion of analog computers, which have been so laboriously constructed in several laboratories.

Since there is a paper by J. Connes covering the subject of computation (Chapter 6), there is no need to consider it in depth here. It might be well to close with the remark that the speed of the FFT is so great that convolution may often be performed faster by transforming into the other domain, multiplying, and transforming back, using the convolution theorem, than by direct convolution. The time required for convolution is $M \times N$, the product of the number of points in the convolvants. If M is significantly larger than $\log_2 N$, then the direct convolution is too inefficient. For example, if phase correction is to be performed using Eq. (1-49), the convolution will take considerably more time than the Fourier transform of the interferogram, if it is done by the conventional method.

1-8 CONCLUSION

Fourier spectroscopy is at the point where it is competitive on a cost basis with any other form of spectroscopy and its very considerable advantages have been proven experimentally. It has been applied in both favorable and unfavorable environments. The inherent simplicity of the apparatus should be appealing to experimentalists, and the delays incurred in computing the spectra can be made negligible. The published theory of Fourier spectroscopy is adequate to cover all cases except the most radically uncompensated interferometers.

To those not using Fourier spectroscopy in their work, we may quote from the Epistle of James, "But be ye doers of the word and not hearers only, deceiving your own selves."

References

1. Connes, J., and Connes, P. (1966) *J. Opt. Soc. Am.* **56**: 896, Connes, J., Connes, P., and Maillard, P. (1969) *Atlas des Spectres Infrarouges de Venus, Mars, Jupiter et Saturne* (Paris, Editions du Centre National de la Recherche Scientifique).

2. Bracewell, R. (1965) *The Fourier Transform and Its Applications* (McGraw-Hill Book Company, New York). This book is the basic reference for the entire mathematical section of this paper.

3. Papoulis, A. (1962) *The Fourier Integral and Its Applications* McGraw-Hill Book Co., New York, p. 269 ff.

4. Vanasse, G. A., and Sakai, H. (1967) in *Progress in Optics*, E. Wolf, ed., North Holland Publishing Company, Amsterdam, Vol. VI, p. 260 ff.

5. Sakai, H., Vanasse, G. A., and Forman, M. (1968) *J. Opt. Soc. Am.* **58**: 84.

6. Jacquinot, P., and Roizen-Dossier, B. (1964) in *Progress in Optics*, E. Wolf, ed., North-Holland Publishing Company, Amsterdam, Vol. III, p. 31 ff.

7. Connes, J., and Nozal, V. (1939) *J. Phys. Radium*, **22**: 339.

8. Chamberlain, J. E., Gibbs, J. E., and Gebbie, H. A. (1969) *Infrared Phys.* **9**: 185.

9. Bell, E. E. (1966) *Infrared Phys.* **6**: 57.

10. Cooley, J. W., et al. (1967) *IEEE Trans. Audio & Electroacoustics*, AU-15: 76.

In lieu of an extensive list of references, the interested reader is referred to Vanasse, G. A., and Sakai (1967) above or to Loewenstein, Ernest V. (1966) *Appl. Opt.* **5**: 845.

Contents

2. Consideration of the Signal-to-Noise Ratio in Fourier Spectroscopy

Hajime Sakai
Air Force Cambridge Research Laboratories
Bedford, Massachusetts

Abstract

Fourier spectroscopy is considered in terms of the signal-to-noise ratio measured. Simple mathematics is used in discussing the basic advantages, multiplex gain, and throughout gain of the technique. Various experimental factors that affect applications are discussed, together with a general expression for the signal-to-noise ratio.

2–1 INTRODUCTION

Fourier spectroscopy provides advantages over the traditional technique because it offers: (1) an increased energy-gathering power through its optical train, and (2) an improvement in the signal-to-noise ratio by means of multiplexing. The advantage realizable from the first factor is theoretically about 100, as shown in comparisons between a Michelson interferometer and a conventional slit spectrometer. The second factor provides a theoretical gain of $\sqrt{N}$ for N spectral elements analyzed. This gain, commonly called the multiplex advantage, has been experimentally confirmed.

This paper presents a general discussion of the advantages of Fourier spectroscopy. Section 2–2 is an elementary explanation of the basic concept: improvement in the signal-to-noise ratio through accumulation of the measured values. The Fourier transformation of the measured interferogram is analyzed as a process that accumulates the spectral signal and the noise in time. Section 2–3 explains the multiplex advantage in Fourier spectroscopy from this point of view. These two sections are designed to furnish an elementary description of measurements in Fourier spectroscopy. Sections 2–4 through 2–7 deal with the technical considerations of the various instrumentation problems. The aim is to derive a reasonable estimate of the signal-to-noise ratio for general measurements performed in the laboratory and in the field. Some mathematical expressions are given without their derivations; more detailed explanations of some of these derivations are given in the appendixes.

2–2 GENERAL CONSIDERATIONS

In Fourier spectroscopy the desired function, the spectrum, is obtained by applying a Fourier transformation to the measured function, the interferogram. The prime factor that determines the quality of the measurements is the ratio of signal to noise in this recovered spectrum, not that in the interferogram.

The following discussion will show that the transformation of either quantity, the signal or the noise, is equivalent to the accumulation. The characteristics of the transformation are such that the longer the signal and noise are recorded during observation of the interferogram, the more the measurement improves. Thus, extending the interferogram measurement in time improves the signal-to-noise ratio of the recovered spectrum, and the improvement is proportional to the square root of time.

The measured value includes some degree of fluctuation, which is generally accepted as the noise; and its root mean square, or standard deviation, is taken as a measure of the noise. For this quantity, say F, a set of values $\{F_1, F_2, F_3, \ldots, F_N\}$ is obtained in N measurements. Figure 2-1 is an example of this set.

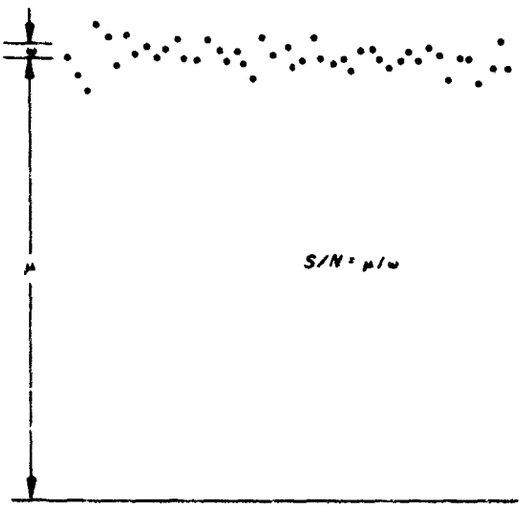


Figure 2-1. Record of Measured Values $\{F_n\}$

From this set we determine the measured value

$$\mu_F = \frac{1}{N}\{F_1+F_2+\cdots+F_N\} = \frac{1}{N}\sum F_n = \langle F_n \rangle, \tag{2-1}$$

and the variance

$$\omega_F^2 = \frac{1}{N}\sum (F_n-\mu_F)^2 = \langle (F_n-\mu_F)^2 \rangle. \tag{2-2}$$

The square root of the variance gives the standard deviation

$$\omega_F = \{\langle (F_n-\mu_F)^2 \rangle\}^{\frac{1}{2}}. \tag{2-3}$$

[In this paper, ω denotes the standard deviation; the more usual symbol σ will be used to denote the optical frequency, cm^{-1}.] Once the mean μ_F and the standard deviation ω_F are found, the true value of F very likely lies in the range between $\mu_F-\omega_F$ and $\mu_F+\omega_F$. The signal-to-noise ratio S/N in this case is defined by

$$S/N = \mu_F/\omega_F. \tag{2-4}$$

Suppose that a new set of (N/m) values $\{G_1, G_2, \ldots, G_{N/m}\}$ is formed by an accumulation of the measured values F_n. Each element of the new set is constructed by summing m elements of the old set. Thus,

$$\begin{aligned} G_1 &= F_1+F_2+\cdots+F_m, \\ G_2 &= F_{m+1}+F_{m+2}+\cdots+F_{2m}, \\ &\vdots \\ G_{N/m} &= F_{N-m+1}+F_{N-m+2}+\cdots+F_N. \end{aligned} \tag{2-5}$$

Figure 2-2 shows the new construction, formed by summation of the successive elements

$$\begin{aligned} G_1 &= F_1+F_2, \\ G_2 &= F_3+F_4, \\ &\vdots \end{aligned}$$

The signal-to-noise ratio has improved somewhat. The actual improvement can be found by computing the mean and the variance. The calculation for the mean shows

$$\begin{aligned} \mu_G = \langle G \rangle &= \frac{1}{N/m}\sum G_n \\ &= \frac{m}{N}\{(F_1+F_2+\cdots+F_m)+(F_{m+1}+\cdots+F_{2m})+ \\ &\quad \cdots+(\cdots+F_N)\} \\ &= m\frac{1}{N}\sum F_n = m\mu_F. \end{aligned} \tag{2-6a}$$

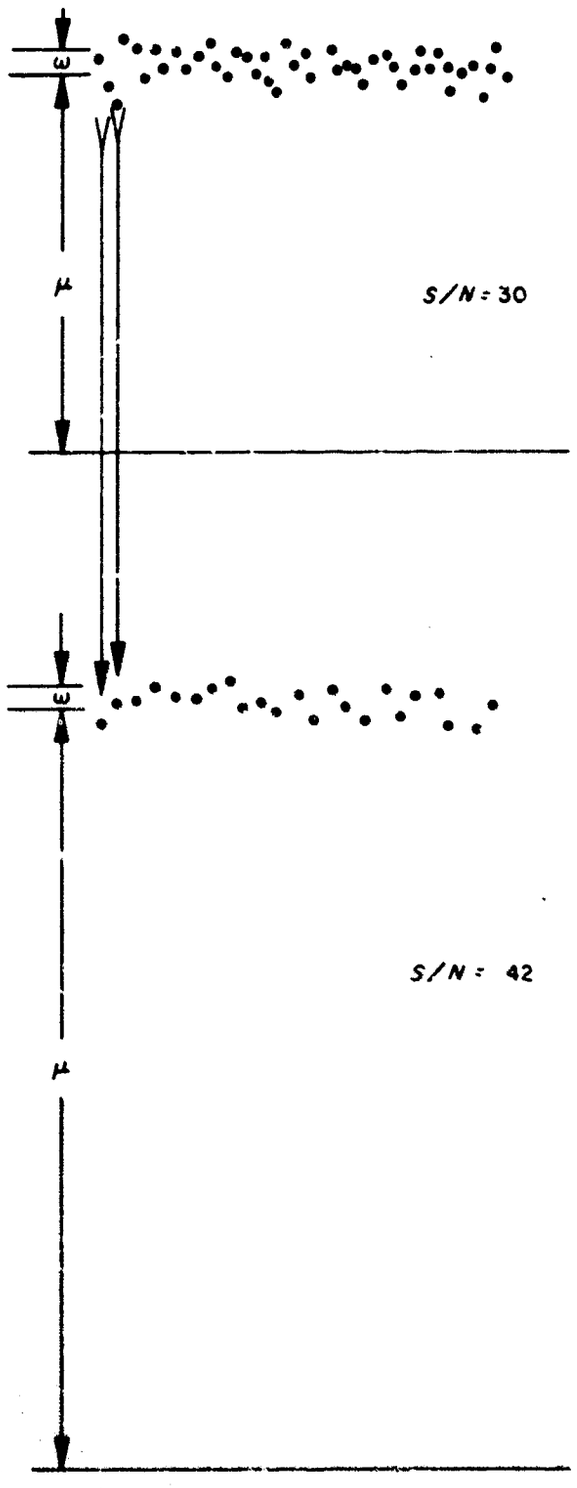


Figure 2-2. Comparison of $\{F_n\}$ and $\{G_n\}$

and the variance is given by

$$\omega_G^2 = \frac{m}{N}\sum (G_n - \mu_G)^2$$

$$= \frac{m}{N}\{(F_1 + F_2 + \cdots + F_m - m\mu_F)^2$$

$$+ \cdots + (F_{N-m+1}$$

$$+ \cdots + F_N - m\mu_F)^2\}$$

$$= \frac{m}{N}\sum_n (F_n - \mu_F)^2 + \frac{m}{N}\sum_n \sum_{n'} (F_n - \mu_F)(F_{n'} - \mu_F)$$

$$= m\frac{1}{N}\sum_n (F_n - \mu_F)^2 + \phi. \qquad (2\text{-}6b)$$

It can be assumed for this case that the second term ϕ is vanishingly small since all values in the measurement are statistically independent. [This assumption will be discussed in calculating the autocorrelation of the noise.] Assuming $\phi = 0$, we can write the variance as

$$\omega_G^2 = m\omega_F^2, \qquad (2\text{-}7)$$

and, consequently, the standard deviation as

$$\omega_G = \sqrt{m}\omega_F. \qquad (2\text{-}8)$$

The signal-to-noise ratio of this new set is then given by

$$(S/N)_G = \sqrt{2}(S/N)_F,$$

and the improvement is $\sqrt{2}$.

In the construction of the new set, m is assumed much smaller than the total number N. Otherwise, the statistical averages taken by using Eqs. (2-5) and (2-6a) are meaningless. As m gets large, it may no longer be possible to perform these statistical averages. Even so, Eqs. (2-6b), (2-7), and (2-8) may be used to

extrapolate an estimate of these parameters. For the case of $m=N$, these estimates are given by

$$\mu = N\mu_F, \tag{2-9}$$

$$\omega^2 = N\omega_F{}^2, \tag{2-10}$$

$$\omega = \sqrt{N}\omega_F. \tag{2-11}$$

We define the autocorrelation as

$$K_F(m) = \langle F_n F_{n+m}\rangle = \frac{1}{N}\sum_n F_n F_{n+m}. \tag{2-12}$$

For convenience ot computation, the subscript of F is assumed cyclic, that is,

$$F_{N+m} = F_m.$$

The variance ω^2 in Eq. (2-10) can now be written in a simplified form by using the autocorrelation:

$$\omega^2 = N\sum_{m=0}^{N-1} K_y(m), \tag{2-13}$$

where the new variable

$$y_n = F_n - \mu_F \tag{2-14}$$

represents the noise fluctuation. The derivation of Eq. (2-13) can be seen in the following mathematical manipulation:

$$\begin{aligned}\omega^2 &= \{F_1+F_2+\cdots+F_N-N\mu_F\}^2\\ &= \{(F_1-\mu_F)+(F_2-\mu_F)+\cdots+(F_N-\mu_F)\}^2\\ &= \sum_n\sum_m (F_n-\mu_F)(F_{n+m}-\mu_F)\\ &= N\sum_m \langle (F_n-\mu_F)(F_{n+m}-\mu_F)\rangle\\ &= N\sum_m K_y(m). \end{aligned}\tag{2-15}$$

If the noise in the measured values is statistically independent, then the noise set $\{y_1, y_2, \ldots, y_N\}$ does not correlate either with sets $\{y_2, y_3, \ldots, y_N, y_1\}$ or $\{y_3, y_4, \ldots, y_N, y_1, y_2\}$ or with any set other than itself. Hence,

$$K_y(m) = \frac{1}{N}\sum y_n y_{n+m} = 0 \quad (m\neq 0) \tag{2-16}$$

if the y_i are statistically independent. Figure 2-3

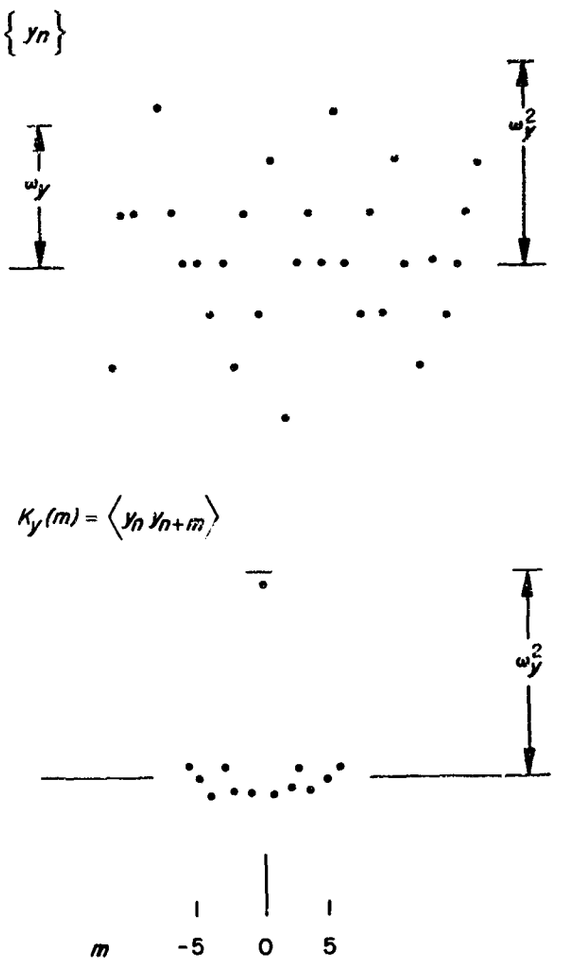


Figure 2-3. Noise Record $\{y_n\}$ and Its Autocorrelation $K_y(m)$

illustrates this property of the noise. The term ϕ in Eq. (2-6*b*) which can be expressed by

$$\phi = N\sum_{m=1}^{N-1} K_y(m),$$

obviously vanishes since all terms in the summation are zero. Combining this with the autocorrelation for $m=0$ yields

$$K_y(0)=\frac{1}{N}\sum_n {y_n}^2={\omega_y}^2. \qquad (2\text{-}17)$$

The autocorrelation can be written in a general form as

$$K_y(m)=\delta_m\frac{1}{N}\sum {y_n}^2={\omega_y}^2\delta_m, \qquad (2\text{-}18)$$

where

$$\delta_m=\begin{cases}1, & \text{for } m=0,\\ 0, & \text{otherwise.}\end{cases}$$

Now suppose that a measurement is made on stationary radiation incoming to a detector. The output of this detector is integrated for 1 sec by means of an electrical filter, and then it is recorded. When this process is repeated N times, a record of N measured values is obtained for N sec (Figure 2-1). From these N values we can determine the statistical parameters mean μ_F, variance ${\omega_F}^2$, and standard deviation ω_F, and estimate the N accumulations by applying Eqs. (2-8), (2-9), and (2-10) to them. The results:

$$\mu=N\mu_F,$$

$${\omega}^2=N{\omega_F}^2,$$

$$\omega=\sqrt{N}\omega_F,$$

are interpreted as those for N sec of integration. Incidentally, the standard deviation ω_F given above for 1 sec of integration is called the noise equivalent power (NEP), the normalized noise figure.

Note in this example that the signal, being stationary in time, consists of a component at zero frequency. The noise, in contrast to the signal, spreads its content over a wide frequency range. An application of the Fourier transformation should separate the signal from the noise content at frequencies other than zero frequency. In Figure 2-4, which shows the transform of $\{F_n\}$, a distinct line is seen at zero frequency, and random fluctuations about zero mean are seen elsewhere. The line at zero frequency obviously corresponds to the stationary radiation signal incoming to the detector. Its height is given by $N\mu_F$, and the standard deviation of the random fluctuation is seen to be

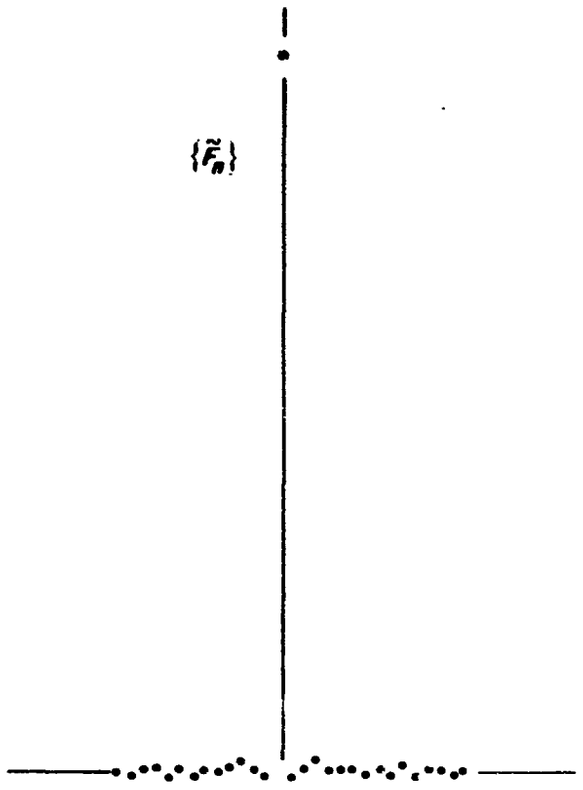


Figure 2-4. Fourier Transform of $\{F_n\}$

$$\omega_{\tilde{F}}=\sqrt{N}\omega_F. \qquad (2\text{-}19)$$

The following computations show that the statistical parameters of the random fluctuation appearing in the transform are given by

$$\mu_{\tilde{F}}=0$$

$${\omega_{\tilde{F}}}^2=N{\omega_{\tilde{F}}}^2$$

$$\omega_{\tilde{F}}=\sqrt{N}\omega_{\tilde{F}},$$

where $\tilde{y}$ is the Fourier transform of y, and

$$\tilde{y}_m = \sum_n y_n \exp\left(i2\pi \frac{nm}{N}\right). \quad (2\text{-}20)$$

Here the arguments, m and n, vary from zero to $N-1$. In the frequency domain, m/N represents values from zero to 1 Hz in increments of $(1/N)$ Hz. The mean of $\tilde{y}$ is obviously zero, since

$$\langle \tilde{y}_m \rangle = \sum \langle y_n \rangle \exp\left(i2\pi \frac{nm}{N}\right). \quad (2\text{-}21)$$

The variance is calculated by

$$\langle \tilde{y}_m \tilde{y}_m^* \rangle$$

$$= \left\langle \sum \sum y_n y_{n'}^* \exp\left(i2\pi \frac{nm}{N}\right) \exp\left(-i2\pi \frac{n'm}{N}\right) \right\rangle$$

$$= \sum \left\{ \sum \langle y_n y_{n+k}^* \rangle \exp\left(i2\pi \frac{km}{N}\right) \right\}$$

$$= \sum \sum K_y(k) \exp\left(i2\pi \frac{km}{N}\right) = \sum \omega_y^2 = N\omega_y^2. \quad (2\text{-}22)$$

The standard deviation of $\tilde{y}$ is then obviously given by

$$\omega_{\tilde{y}} = \sqrt{N}\omega_y. \quad (2\text{-}23)$$

Identical results are now found for both standard deviations, one for the transformed and the other for the accumulated. That the agreement is not accidental can be seen from the calculations given above. It is indicated in Eq. (2-20), which defines the transformation of the noise. This equation can be interpreted as an accumulation of the original y_n multiplied by the exponential weighting factor $\exp(i2\pi nm/N)$. In the complex domain, these weighting factors may have a different phase but their magnitude never changes from unity.

The signal-to-noise ratio of the Fourier transform is given by the signal content divided by the standard deviation of the transformed noise. Hence

$$(S/N)_{\tilde{F}} = \mu_{\tilde{F}}/\omega_{\tilde{y}} = \sqrt{N}\mu_F/\omega_F = \sqrt{N}(S/N)_F, \quad (2\text{-}24)$$

which is equal to that of the accumulated. Thus, the improvement brought through the transformation is $\sqrt{N}$, because N values measured during the observation over N sec in time are accumulated in the transform. In other words, the improvement is achieved during an observation of N sec rather than of only 1 sec. The Fourier transformation accomplishes this accumulation automatically.

2-3 MULTIPLEX ADVANTAGE

So far it has been shown that the improvement in the signal-to-noise ratio is $\sqrt{N}$ if the measurement time is extended by N. Suppose that N spectral elements are measured by two different methods, one sequential and the other simultaneous. Both methods will yield the same value if they take an equal time to measure each element. If we compare the total time elapsed for each method, we find that the sequential scheme takes N times longer than the simultaneous scheme. Suppose now that the elapsed times are made equal. Under this condition, the observation of each element will take N times longer in the simultaneous scheme than in the sequential scheme. Thus, as has been shown, the signal-to-noise ratio in the simultaneous measurement is $\sqrt{N}$ times higher than in the sequential measurement. This is the basis of the multiplex advantage that can be realized by applying the Fourier spectroscopy technique. Examples of these two methods are found in two schemes for spectral measurements, one a sequential scheme based on the classical spectrometer used with one detector, and the other a multichannel scheme based on the spectrometer used with N detectors.

Figure 2-5 illustrates the scheme in which each element is sequentially observed for one unit of time. An element at $\sigma = \sigma_n$ is observed during $(n-1) \leq t \leq n$, and then the spectrometer is scanned to the next position $\sigma = \sigma_{n+1}$. The spectrum is presented in the output signal as a function of time.

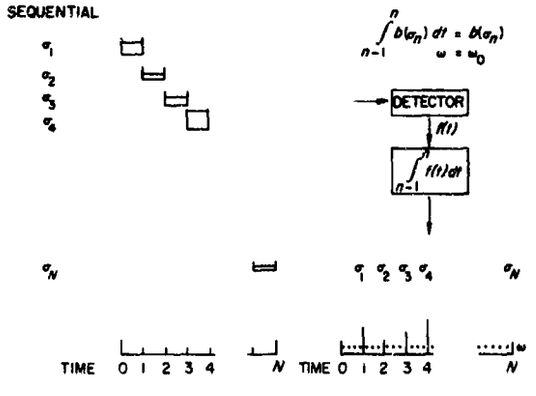


Figure 2-5. Sequential Measurement Scheme

The multichannel scheme illustrated in Figure 2-6

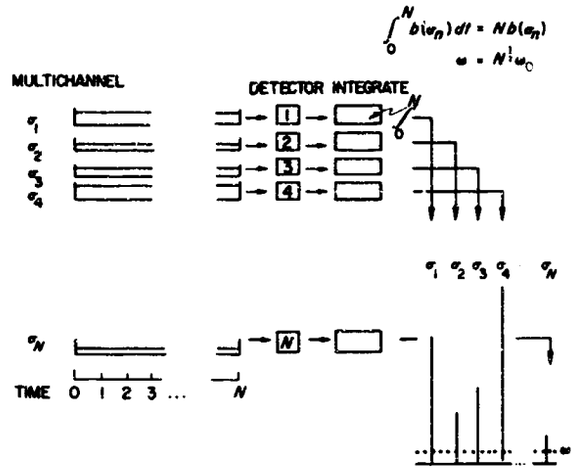


Figure 2-6. Multichannel Measurement Scheme

uses N detectors having identical characteristics. Each detector observes each spectral element for N units of time during $0 \leq t \leq N$, at the end producing the spectral signal for the element.

If the power incoming to the detector is $b(\sigma)$ at the spectral element σ, the signal for the sequential scheme is given by

$$\int_{n-1}^{n} b(\sigma)\, dt = b(\sigma). \tag{2-25}$$

The noise is characterized by ω. The signal for the multichannel scheme is

$$\int_{0}^{N} b(\sigma)\, dt = Nb(\sigma), \tag{2-26}$$

and the noise is then given by $\sqrt{N}\omega$. The signal-to-noise ratios are therefore $b(\sigma)/\omega$ for the sequential, and $\sqrt{N}b(\sigma)/\omega$ for the multichannel, schemes. The multichannel measurement thus shows an improvement of $\sqrt{N}$ over the sequential measurement.

The multiplex scheme used in Fourier spectroscopy is illustrated in Figure 2-7. A single detector is used.

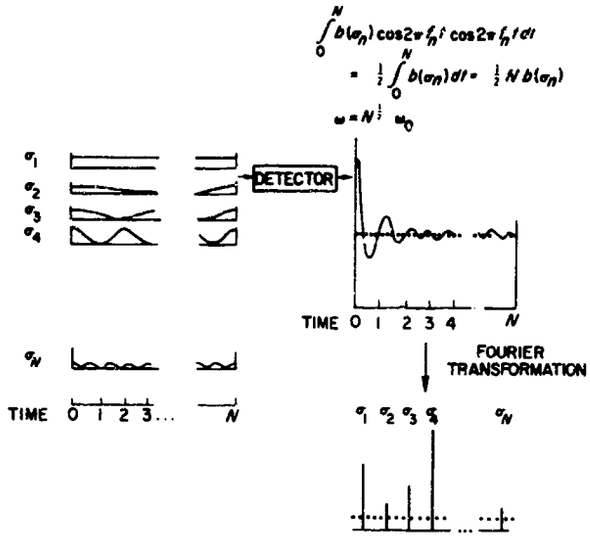


Figure 2-7. Fourier Spectroscopy Measurement Scheme

Each spectral element is coded with a cosine function $\cos 2\pi f_s t$. The temporal frequency is given by

$$f_s = \sigma v \tag{2-27}$$

where v is the drive speed of the interferometer mirror. These coded signals are simultaneously observed for N units of time during $0 \leq t \leq N$. Since only half of the incoming energy is available for detection with this coding because the other half must go back to the source, the signal at the spectral element σ is given by

$$\int_{0}^{N} \tfrac{1}{2} b(\sigma)\, dt = \frac{N}{2}\, b(\sigma). \tag{2-28}$$

The noise is $\sqrt{N}\omega$, and so the signal-to-noise ratio is given by $\frac{1}{2}\sqrt{N}b(\sigma)/\omega$. The improvement in the signal-to-noise ratio over the result obtained in the sequential scheme is $\frac{1}{2}\sqrt{N}$. In most treatments, the factor $\frac{1}{2}$ is dropped on the assumption that equal energy is available for detection. Measurements by the Fourier technique show the multiplex gain of $\sqrt{N}$ over measurements by the sequential technique only if the detectable energy is equal. A comparison of the two modes in the Fourier technique shows the $\sqrt{N}$ improvement in results obtained with the aperiodic over those obtained with the periodic.[1]

2-4 SIGNAL DETECTION BY INTERFEROMETER

It has been found that the multiplex method can improve the signal-to-noise ratio over that yielded by the sequential method by the square root of the number of elements in the spectrum. This multiplex gain is obtained under two conditions: (1) that both systems take the same total observation time; (2) that both systems observe the same radiative energy in a unit time for each element. A typical sequential technique uses a classical slit spectrometer whereas the Fourier technique depends mainly on a Michelson interferometer. A comparison of these two instruments reveals a rather large difference in energy-gathering efficiency.[2,3]

Figure 2-8 is a schematic diagram of the Michelson

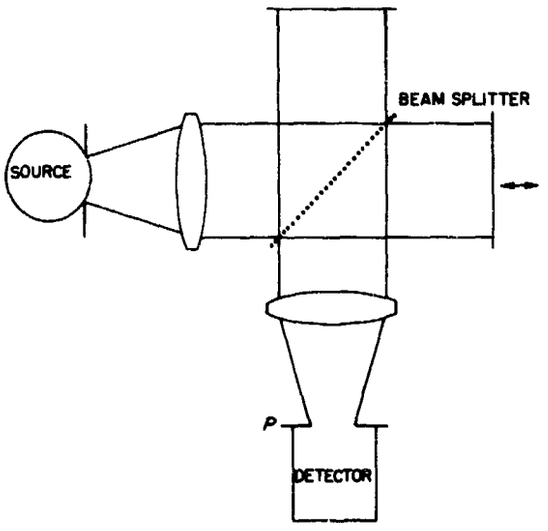


Figure 2-8. Schematic Diagram of Fourier Spectroscopy Measurement System With a Michelson Interferometer

interferometer. Radiation energy from the source is collimated and fed into the interferometer. Except on rare occasion, an uncollimated input is not used because this would lower the energy-gathering efficiency.[4]

The output from the interferometer is focused on a plane P to form the Haidinger fringes. An aperture stop is usually placed in this plane and the radiation through this opening reaches the detector. The aperture stop located in front at the source can be eliminated from consideration since it is the image of the detector aperture. The source emits radiation through a real or an imaginary opening. If a unit area of the source emits a radiative energy $b(\sigma)$ at frequency σ over a unit solid angle for a unit time and unit frequency interval, the interferometer receives radiation energy at a rate of $b(\sigma)(\delta\sigma)[A\Omega]$/sec, where A is the area of the source opening, Ω is the solid angle subtended by the entrance aperture, and $\delta\sigma$ is the spectral width of the radiation. The product $A\Omega$, called the optical throughput or the *étendue*, remains invariant throughout the optical system and hence equal to that defined by the image on plane P.

The radiation that enters the interferometer is divided into two wavefronts by the beam splitter and is then recombined to form the interferogram. If the area of the detector is reasonably small, the energy E going toward the detector is given by

$$E = 2\eta b(\sigma)(\delta\sigma)[A\Omega]\{1+\cos 2\pi\sigma x\}/\text{sec}, \quad (2\text{-}29)$$

where η is the efficiency of the beam splitter (normalized to 0.25 for the maximum value) and x is the path difference between the two interfering beams (see Appendix 2-A.1). On plane P, interference forms the system of Haidinger fringes shown in Figure 2-9. As

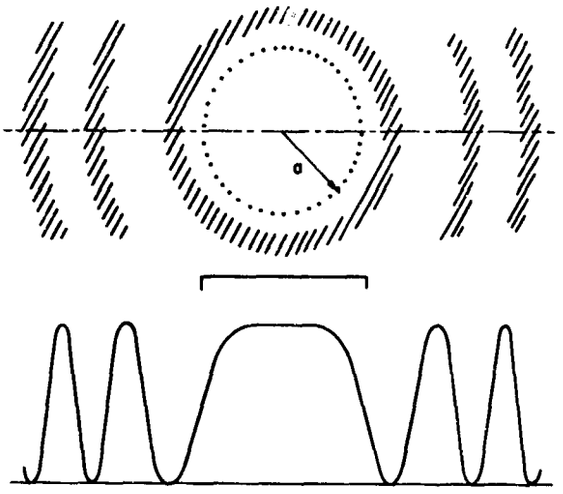


Figure 2-9. Radius a of Detector Aperture

the interference path difference x increases, these rings expand and a new ring keeps appearing at the center. Equation (2-29) describes the energy measured at the center of the ring system. From the figure it immediately becomes evident that the detector aperture cannot be expanded too much beyond the size indicated by the dotted circle. Since the ring system of a higher frequency gets more compact than that of a lower one, the largest permissible detector aperture is determined by the highest frequency present in the spectral bandwith. The angular radius of the aperture is given by

$$\alpha = \{1/(2\sigma X)\}^{\frac{1}{2}}, \quad (2\text{-}30)$$

where X is the maximum path difference. It can also be expressed very simply in terms of the maximum resolving power R as

$$\alpha = \{1/R\}^{\frac{1}{2}}. \tag{2-31}$$

For example, to obtain a resolving power of 10^6, the angular diameter 2α of the detector should be set at

$$2\alpha = 2(10^{-6})^{\frac{1}{2}} = 2\times 10^{-3} \approx 7 \text{ min.}$$

The solid angle corresponding to this is given by

$$\Omega = \pi\alpha^2. \tag{2-32}$$

Thus, the optical throughput for an interferometer is given by

$$A\Omega = \pi A/R, \tag{2-33}$$

whereas for a slit spectrometer it is given approximately by[5]

$$A\Omega = (A/R)(l/f), \tag{2-34}$$

where l is the height of the slit and f is the focal length of the telescope lens or mirror. Since the factor (l/f) is unlikely to exceed 1/30, it is possible to make the optical throughput of an interferometer enormously larger than that of a slit spectrometer of comparable spectral resolution.

A practical beam splitter film for the infrared region is a dielectric thin film. Its actual efficiency can be predicted with reasonable accuracy[6-8] by a straightforward theoretical calculation.[9] Figure 2-10 shows

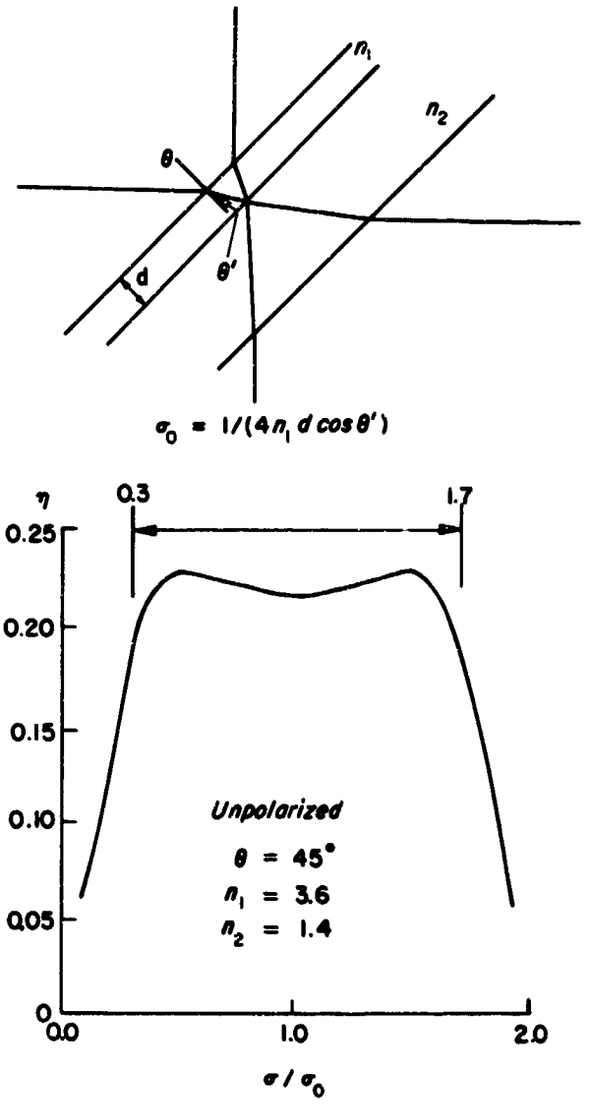


Figure 2-10. Efficiency of Beam Splitter Film

an efficiency curve for a typical beam splitter. The normalized frequency σ_0 used in its frequency scale is given by

$$\sigma_0 = 1/(4nd\cos\theta'), \tag{2-35}$$

where n is the refractive index of the beam splitter film, d is the thickness of the film, and θ' is the refraction angle inside the film. In practice, an efficiency close to the maximum value 0.25 is obtainable by using a dielectric material of high refractive index.[6] This accounts for the wide use of Ge ($n = 4.0$) and Si ($n = 3.4$), whose range of practical frequencies lies between $0.3\sigma_0$ and $1.7\sigma_0$.

With all these factors considered, the energy available for spectral recovery is

$$E = \eta_1' \pi A b(\sigma) \{\delta\sigma\} T/R, \qquad (2\text{-}36)$$

where T is the time interval available for measuring the entire interferogram. The efficiency η_1' is that of the interferometer, which should be very close to the efficiency of the beam splitter.

2-5 TYPES OF NOISE

Spectral measurements obtained by applying techniques of Fourier spectroscopy may include fluctuations due to detector, photon, scintillation, or digitizing noise. This section covers the individual characteristics of these four types of noise and their various effects on the measurements.

The detector noise determines the accuracy of most measurement practices, and the photon noise from the source determines the ultimate accuracy of the measurements. The effects of scintillation and digitizing noise must be suppressed in the compensation schemes.

2-5.1 Detector Noise

The mechanism that generates noise in the detector is generally insensitive to the signal level of the radiation incoming to the detector, and consequently the amount of noise remains unchanged (Figure 2-11).

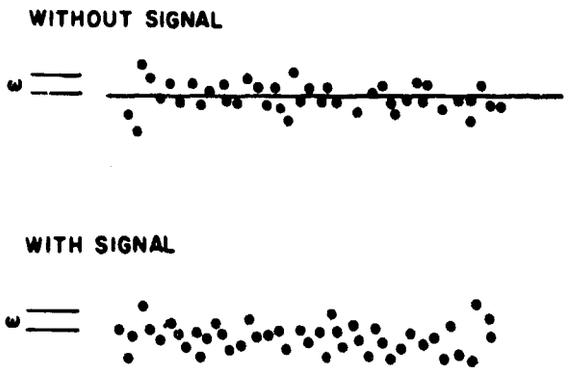


Figure 2-11. Detector Noise

In Fourier spectroscopy, the radiation incoming to the detector has a wide spectral range and its intensity level may be very high. This is quite contrary to radiation detection conditions in the case of classical spectroscopy. The detector noise remains the same in both cases even though there may be a vast difference in the intensity levels. The multiplex advantage of Fourier spectroscopy over the classical scheme is realized if the detector noise dominates other noise fluctuations.

The noise figure of a detector is usually characterized in terms of the normalized standard deviation; it is determined from its observed noise record integrated for a time constant of 1 sec. Called the noise equivalent power (NEP) of the detector, this is expressed in units (watts) of the incoming radiation. Another noise figure is defined by

$$D^* = \sqrt{A}/(NEP), \qquad (2\text{-}37)$$

where A is the sensitive area of the detector element.

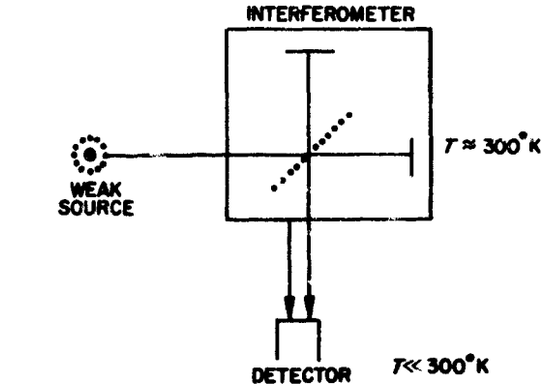
2-5.2 Photon Noise

The stream of photons from a thermal source exhibits an inherent fluctuation proportional to the square root of its average intensity.[10] In the infrared region, this photon noise is insignificant if the spectral measurement is referred to the source (see Appendix 2-A.2) but creates a serious problem if the measurement includes the background.

In Fourier spectroscopy, a warm interferometer may emit intense thermal background radiation toward a detector that operates at a low temperature. In such case the detector observes a relatively weak source signal and a strong background radiation (Figure 2-12). The fluctuation in the background may become dominant in the measurement. The situation is somewhat parallel to the conventional case where a warm spectrometer generates the thermal background.

In either case—Fourier spectroscopy or the conventional scheme—the signal measurement may be affected by the intensity fluctuation in the background, not by the detector noise. By extending the observation of the signal in time the measurement can be improved. The improvement achieved by means of the Fourier technique is greater than that obtained by the conventional method because the observed signal is multiplexed. If the level of the background radiation is equal in both cases, the Fourier spectroscopy technique can be expected to yield an improvement of $\sqrt{N}$ over the results given by the conventional method.

The photon noise in the source signal may become noticeable in the spectral region of higher optical frequency, for which detectors having an extremely high sensitivity are available. If the photon noise from the source itself dominates the detector noise, the multiplex advantage is not realized.



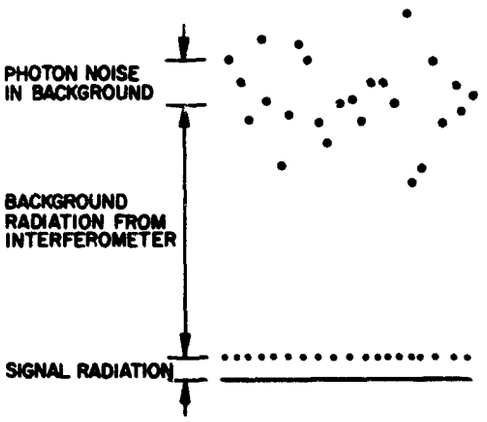


Figure 2-12. Background Photon Noise From a Warm Interferometer

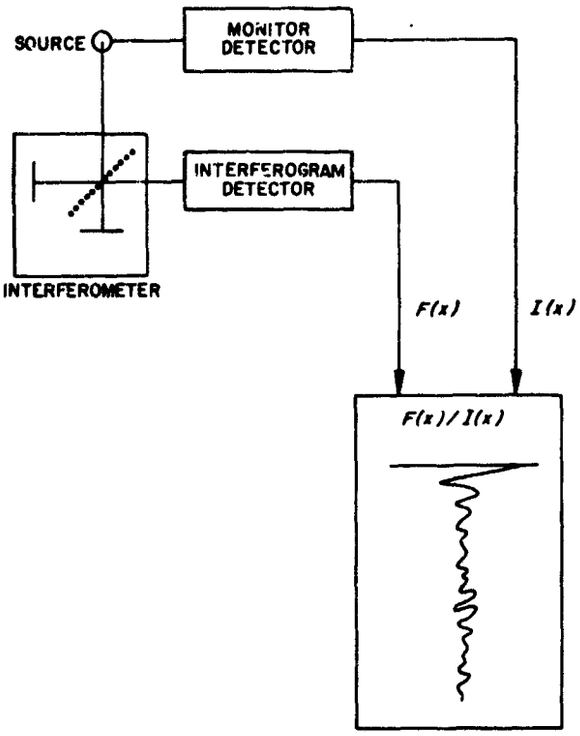


Figure 2-13. Ratio-Recording Scheme

2-5.3 Scintillation Noise

Scintillation noise generally refers to a slow drift in the level of intensity observable at the interferometer input. Its magnitude is linearly proportional to signal intensity. The fluctuation may cause serious degradation of the spectral recovery.[11-13] The several techniques applied to minimize the effect of this fluctuation include:

a) ratio-recording[11-13]

b) rapid scanning[13]

c) internal modulation[12,13]

d) beam-switching

The first three schemes are applicable to astronomical observations as well as to laboratory measurements. The fourth, discussed in Chapter 4 of this collection (Double-Beaming in Fourier Spectroscopy, by J. Dowling), applies only to laboratory measurements. Schemes *b* and *c* are described in Section 2-6, where various experimental arrangements are discussed.

The scheme of ratio-recording shown in Figure 2-13 uses an additional detector for monitoring the input intensity level. The interferogram signal is then normalized by this monitored intensity.

2-5.4 Digitizing Noise

In practice, a digital computer is used to perform a numerical Fourier transformation on a recorded interferogram, which in this case consists of a set of numbers obtained from the raw signal as an output voltage of the detector amplifier. Various types of analog-to-digital (A-D) converters are used for this purpose. Some are very sophisticated electronic instruments. Others may consist of a primitive device in which a human operator reads an interferogram trace recorded on chart paper and converts it into a set of numbers. Regardless of the method, a small but finite range of a conversion aperture Δ exists. The output from the A-D converter corresponds to any value of the detector output in the range $I_0 \pm \frac{1}{2}\Delta$; and the error takes any value between $\pm\frac{1}{2}\Delta$. If values of this error are displayed along the path difference x, the result is a random function of x (Figure 2-14) called the digitizing noise. Its standard deviation is given by

$$\omega = \left\{\frac{1}{\Delta}\int_{-\Delta/2}^{\Delta/2} x^2\,dx\right\}^{\frac{1}{2}} = \frac{\Delta}{2\sqrt{3}} \approx 0.29\Delta.$$

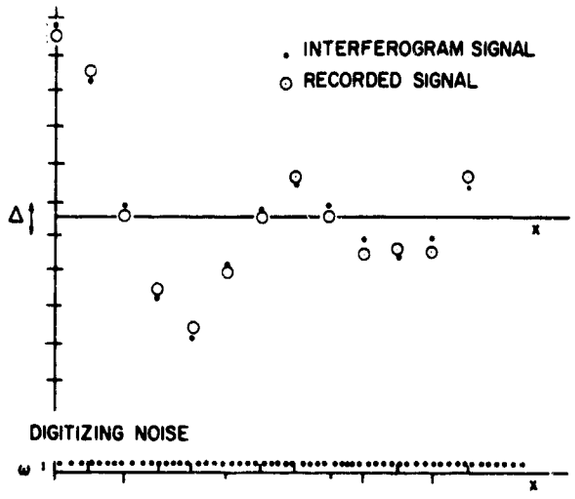


Figure 2-14. Digitizing Noise

This noise transforms into a random fluctuation in the spectral domain as described in Section 2-1.

Serious difficulty may be encountered in the measurement when the digitizing noise exceeds the detector noise. A practical solution to this problem is to vary the gain of the detector amplifier so that the detector noise dominates the digitizing noise over most of the recorded interferogram. It is generally safe to assume that modulation in the interferogram decreases once the path difference exceeds a certain value. In a limited range around zero path difference, the interferogram measurement is made with the gain in the detector amplifier reduced to accommodate the increased modulation and high level of digitizing error. Outside this range, where a reduced modulation can be anticipated, the amplifier gain is increased to let the detector noise exceed the digitizing noise. In computing the Fourier transformation, the values that are recorded with different gains are normalized according to the gain setting that is used. For most of the measurement the interferogram contains detector noise that exceeds the digitizing noise.

With more sophisticated electronics, the dynamic range of the A-D conversion can be automatically changed according to the level of the detector output (automatic ranging). The error generated in this case is similar to the round-off error in floating-point arithmetic. Several workers have attempted to analyze this error.[11-15]

2-6 EFFECT OF ERRORS IN SAMPLING INTERVALS

The interferogram is a function of the path difference x. The function consists of a finite number of discrete values sampled at equispaced distances on the path-difference scale. The path difference is increased by a given distance as the recording of the interferogram proceeds from the nth to the $(n+1)$th sample. The intervals between successive sampling positions are selected to suit the tolerance required by the experiment. Excessively large errors in the sampling intervals will create unwanted effects in the recovered spectrum, degradation of the spectral resolution, and additional noise fluctuation. Similar effects occur in conventional spectroscopy when grating-ruling errors are present, although not as pronounced as those occurring in Fourier spectroscopy.[16]

Sinusoidal ruling errors on a diffraction grating produce the well-known phenomenon of the Rowland ghost. In Fourier spectroscopy, sinusoidal errors in the sampling intervals produce the same effect. A monochromatic interferogram sampled with the sinusoidal errors yields, through cosine transformation, a line and two satellites. The height of these satellites is given by

$$\beta = \tfrac{1}{4}\theta\sigma_0, \tag{2-38}$$

where β is normalized with respect to the height of the center line, θ is the amplitude of the sinusoidal sampling errors, and σ_0 is the frequency of the original monochromatic line. The satellites are located at $\sigma_0 \pm \varphi$ where $1/\varphi$ is the periodicity of the errors. (See Appendix 2-A.3 for the derivation of these results.)

In most interferogram measurements, the errors in the sampling intervals are best represented by a random function rather than a sinusoidal function. Then the recovered spectrum may contain more satellites, each pair of which corresponds to sinusoidal errors having a particular periodicity. The amplitude of the sinusoids is obviously a random function given by the Fourier transform of the sampling errors. Based on Eq. (2-8), the standard deviation of the transform θ is calculated as

$$\omega_\theta = \sqrt{R}\omega_x, \tag{2-39}$$

where R is the number of the interferogram samples and ω_x is the standard deviation of the sampling errors. The term R is in most cases equal to the resolving power. Once the standard deviation of the sinusoidal amplitude θ is known, the standard deviation of the height of those satellites that obviously have the appearance of random noise can be calculated from Eq. (2-38) as

$$\omega_\beta = [\langle \beta^2 \rangle]^{\frac{1}{2}} = \tfrac{1}{4}\omega_x\sigma_0\sqrt{R}. \tag{2-40}$$

It is now seen that the random errors in the sampling intervals generate many satellites that have the appearance of random noise. This effect can be considered as a degradation of the spectral resolution and/or the addition of random noise. The calculation of a tolerance for the sampling errors ω_x can be made by setting an upper bound on ω_x. (See Appendix 2-A.4 for the mathematical expression of the tolerance.) In Figure 2-15, tolerances ω_x are shown for

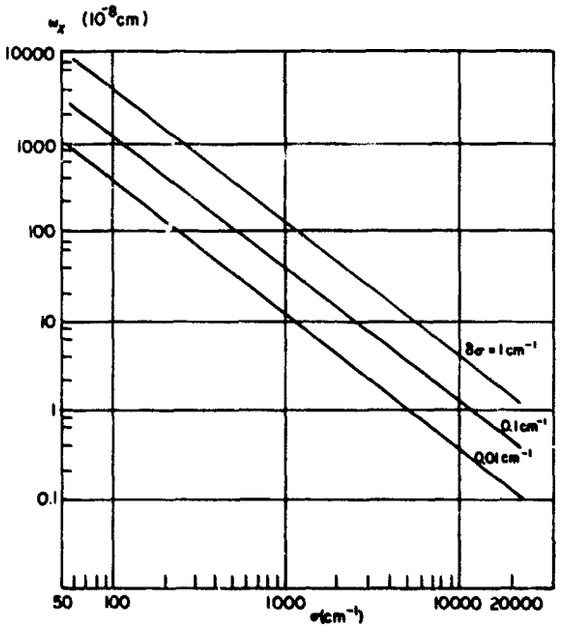


Figure 2-15. Tolerance for Sampling Error, $\omega_\beta = 0.01$ (see Section 2-6 and Appendix 2-A.4)

given values of the resolution $\delta\sigma$. The permissible upper bound of the additional noise ω_β is assumed to be 0.01 in these curves.

2-7 EXPERIMENTAL TECHNIQUES

Various factors that affect the signal-to-noise ratio have been treated in Sections 2-5 and 2-6. The results obtained will now be applied in the analysis of some experimental schemes. In a broad sense, these schemes fall into two groups according to their characteristic manner of driving the interferometer mirror: (1) the continuous-drive method and (2) the step-and-integrate method. Figure 2-16 illustrates their patterns. The discussion deals with: (1) efficiency of use of the observation time, (2) compensation for scintillation noise, (3) sampling error.

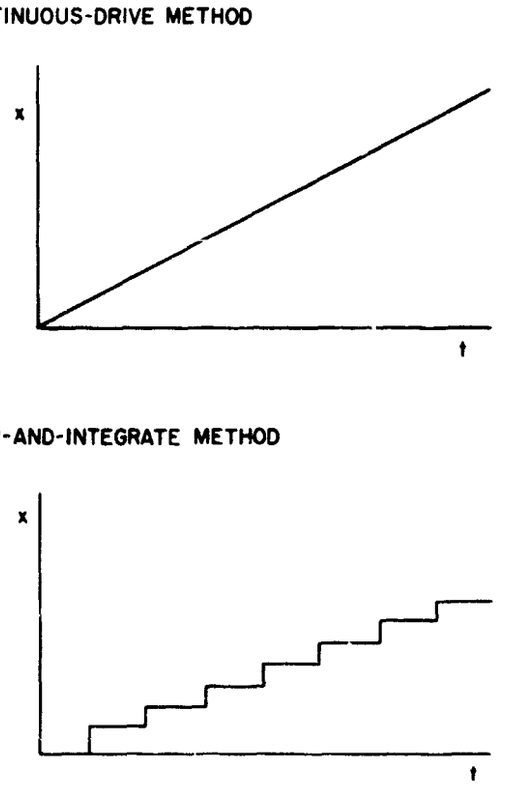


Figure 2-16. Temporal Variation of Path Difference x in Continuous-Drive Method and in Step-and-Integrate Method

2-7.1 Continuous-Drive Method

For the continuous-drive method to be successful,[17,18] the interferometer mirror must be smoothly driven at a uniform speed (Figure 2-16). The simplest experimental scheme is illustrated in Figure 2-17. A linear change of the path difference with time is presumed, and the interferogram is sampled at equal time intervals. The positions of the samplings are equispaced on the path difference scale since the drive speed is uniform. Considering the mechanical tolerance of available drive systems, the scheme is suitable for work requiring a relatively low resolution ($\mathcal{R} \approx 100$) in the far infrared region.

An electrical filter circuit is normally attached to the detector amplifier to effect temporal intergration of the interferogram signal between samplings. Otherwise, the time intervals between the samplings would be lost from the observation. The filter, having a time constant τ, yields the value of the interferogram signal integrated over τ at every sampling position. With N interferogram samplings, it produces spectral signal observations over the time $N\tau$, the total time being T. The observation efficiency η is then given by

$$\eta = N\tau / T. \qquad (2\text{-}41)$$

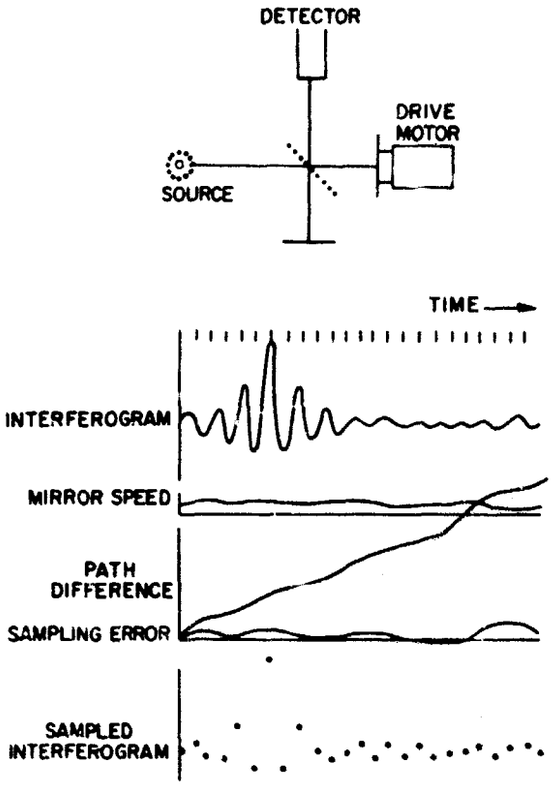


Figure 2-17. Continuous Drive Method Without Fringe-Monitoring Device

or it can be expressed by

$$\eta = \tau/(\Delta t), \tag{2-42}$$

using the time interval Δt between successive samplings. A simple RC circuit can be used as the electrical filter.[14] Caution must be exercised in selecting the time constant τ with respect to the sampling interval Δt. This point is discussed in Appendix 2-A.5.

The scheme illustrated in Figure 2-18 improves the

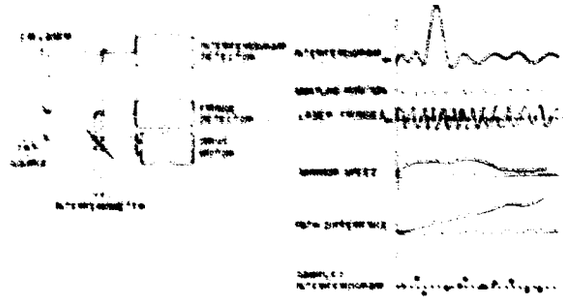


Figure 2-18. Continuous-Drive Method With Fringe-Monitoring Device

positioning accuracy of the interferogram sampling. Monochromatic radiation from a He-Ne cw laser produces the interference fringes for monitoring the path difference x. Here, too, an electrical filter improves the observation efficiency. If the mean time interval between successive sampling positions is Δt, the observation efficiency is again given by Eq. 2-42.

A difficulty arises if the drive speed of the interferometer mirror fluctuates. Figure 2-19 illustrates the

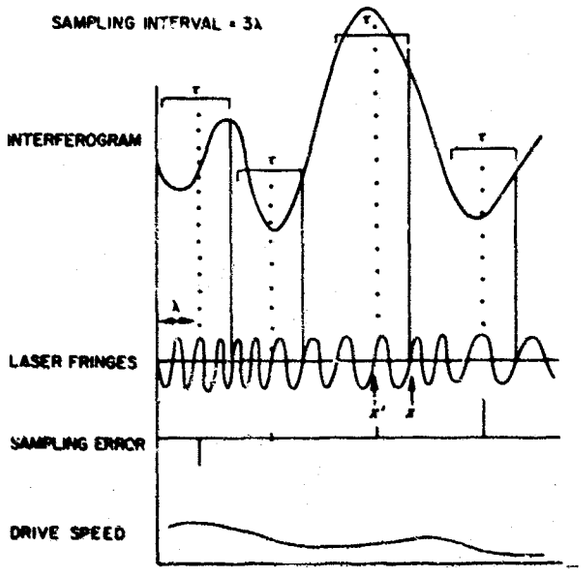


Figure 2-19. Sampling Errors Caused by Mirror Speed Fluctuation

problem: the effective sampling positions x' do not coincide with the positions x of the laser command, unless the filter time constant τ is an infinitesimal value. The midpoint between the beginning and end of the filter integration is the effective sampling position x'. The distance between these two positions is therefore given by

$$x' - x = \tfrac{1}{2}\tau v. \tag{2-43}$$

When the speed varies by Δv from its mean speed v, the error introduced in the sampling position is given by

$$\Delta x = \tfrac{1}{2}\tau\,\Delta v. \tag{2-44}$$

The undesirable effect discussed in Section 2-6 appears in the spectrum if the sampling error expressed in Eq. (2-44) exceeds the tolerance. For a given value of $\Delta\nu$, the filter time constant τ must be reduced in order to suppress the error to a figure within the tolerance. In turn, the measurement shows a decreased observation efficiency η. Table 2-1

Table 2-1. Observation Efficiency of Continuous-Drive Method

Speed Fluctuation $\Delta\nu/\nu = 0.10$ Additional Fluctuation $\omega_\beta = 0.01$ Sampling Distance $\chi = 0.6328 \times 10^{-4}$ cm		
Frequency (σ cm^{-1})	Resolution (σ cm^{-1})	Efficiency (η)
1000	1.0	0.40
1000	0.1	0.15
5000	10.0	0.15
5000	1.0	0.05
5000	0.1	0.02
10000	10.0	0.05
10000	1.0	0.02

lists the values of η acceptable for a speed fluctuation $\Delta\nu/\nu = 0.10$ and for $\omega_\beta = 0.01$. The optimum efficiency available for a given speed fluctuation $\Delta\nu/\nu$ is calculated as

$$\eta = \frac{0.16}{\sqrt{R}} \frac{\nu}{\Delta\nu}. \qquad (2\text{-}45)$$

[See Appendix 2-A.6 for the derivation of Eq. (2-45).]

It is evident from Eq. (2-45) and Table 2-1 that a straightforward application of this scheme to high-resolution work produces a rather poor observation efficiency; the signal-to-noise ratio in the recovered spectrum may not approach the theoretical value of the multiplex advantage. (Several attempts have been made to overcome this problem. For an example, see Sakai and Murphy.[13]) Even so, a large energy-gathering power is available through the interferometric technique. It uses a relatively simple drive mechanism that can deliver a fast mirror speed, an essential feature for rapid scanning.[13] Since the scintillation frequency ranges below 1 Hz, the intensity remains unchanged in a time period that is much shorter than 1 sec. The rapid scanning scheme records an entire interferogram signal within this period. Continuous repetition of this recording process eventually builds up the observation necessary for the desired measurement.

2-7.2 Step-and-Integrate Method

In the step-and-integrate method, the interferometer is driven in a staircase pattern, as shown in Figure 2-16. The interferogram signal is produced by integrating the detector output signal during the time interval that the mirror is held stationary. The time period during which the mirror is stepped from one sampling position to the next is lost time that does not contribute to the observation; keeping these periods short by comparison with the integration periods yields a high observation efficiency. The gain in signal-to-noise ratio obtained is close to the theoretical value of the multiplex gain.[11,19] A simple scheme for this method depends on a drive mechanism consisting of a machine screw and a stepping motor to move the mirror to the sampling positions.[19] Its success is determined by the accuracy of these components since they relate to the sampling error tolerance. It has been shown that the system is satisfactory for far-infrared work.

In the system shown in Figure 2-20, an improved

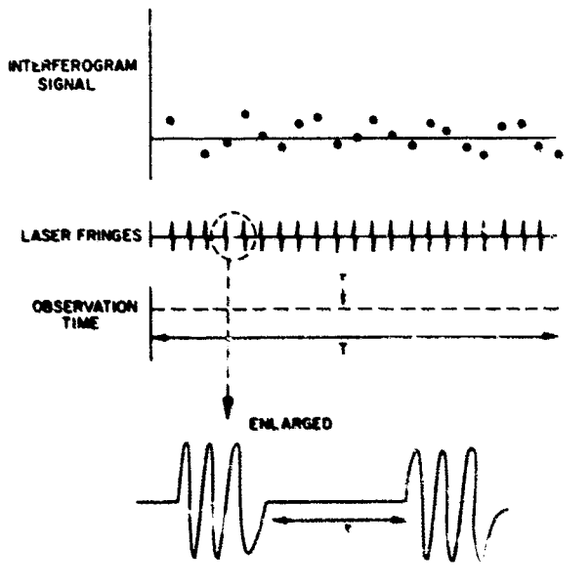


Figure 2-20. Step-and-Integrate Method

positioning accuracy for the sampling is provided by a servomechanism control system that incorporates the monochromatic interference fringe signal. Radiation from a He-Ne cw laser is again commonly used to produce the fringes. Connes et al[11,12,20] have used this means of obtaining spectral measurements of extremely high resolution for laboratory work as well as for astronomical observations.

Since the drive mechanism of this scheme can operate only at a moderate speed, the method cannot accommodate the rapid-scanning scheme. Scintilla-

tion noise is compensated for by using either the ratio-recording[11,20] or the internal-modulation[12,13] schemes. When the intensity drift is due to atmospheric scintillation, the signal intensity observed by the monitor may not correlate with that in the interferogram, as in the situation illustrated in Figure 2-21.

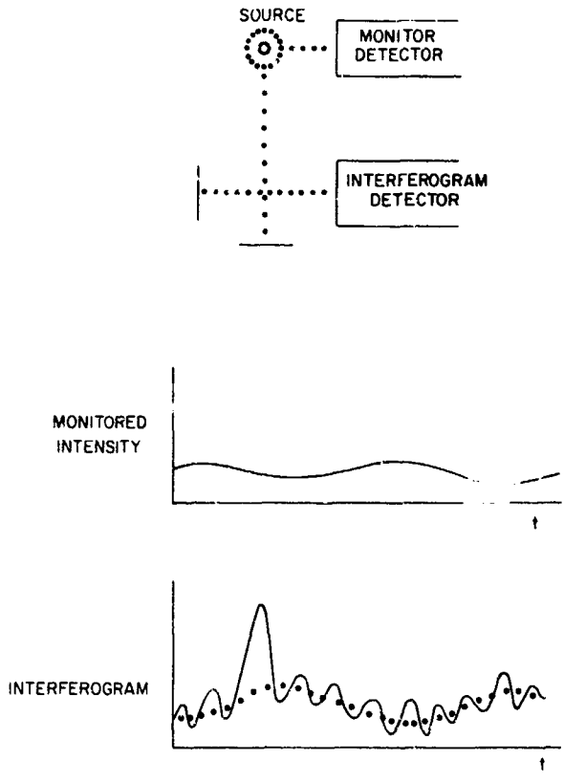


Figure 2-21. Example Showing Correlation Between Intensity-Monitoring Signal and Interferogram Signal

The ratio-recording scheme fails to work under such a condition. The internal-modulation scheme uses the reference intensity internally generated. The reference signal is provided by the interferogram signal of another path difference. In practice the measurement is made at two sampling positions separated by $\pm\Delta$ (see Figure 2-22). Reference is accomplished by rapidly subtracting one signal from another. The signal thus obtained is given by

$$F'(x) = F(x+\Delta) - F(x-\Delta), \qquad (2\text{-}46)$$

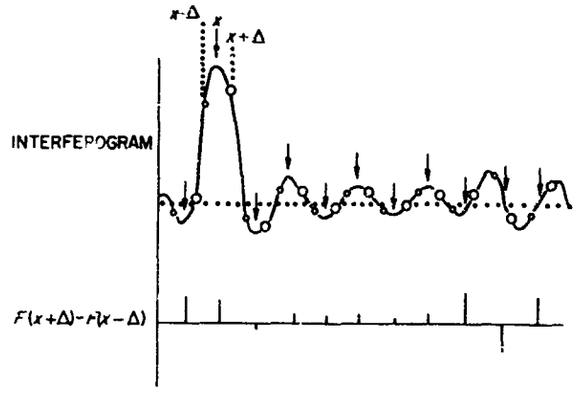


Figure 2-22. Internal-Modulation Scheme

and its Fourier transformation by

$$\int \{F(x+\Delta) - F(x-\Delta)\} e^{-2i\pi\sigma x}\, dx = 2i(\sin 2\pi\sigma\Delta) B(\sigma). \qquad (2\text{-}47)$$

The spectrum is recovered as a product of a sine function and the original spectrum $B(\sigma)$. The practical maximum for the separation Δ is then given by

$$\Delta = 1/(2\sigma_M), \qquad (2\text{-}48)$$

where σ_M is the highest frequency in the bandwidth. Now we can see the elimination of the baseline modulation. This indicates that the recovered spectrum at zero frequency is zero. The atmospheric scintillation is compensated for by suppressing the signal content in the frequency region below about 1 Hz, the highest scintillation frequency.

2-8 SUMMARY

Spectral measurements obtained by applying techniques of Fourier spectroscopy may include fluctuations caused by four types of noise: detector, background photon, scintillation, and digitizing. The effect of the first two can be reduced by an extension of the observation time. If these two are dominant, the signal-to-noise ratio is given by $S/N \propto \sqrt{T}/(NEP)$ —where $\sqrt{T}$ is the incoming energy to the detector per second—and a full realization of the multiplex gain is possible.

The effect of scintillation noise must be reduced by means of some compensation scheme.

The digitizing noise must be reduced to the level of the detector noise by adjusting the dynamic range during the interferogram measurement.

The rate of energy incoming to the detector through a given interferometer is given by

$$B = \pi \eta b(\sigma)\ \delta\sigma A/R.$$

The sampling error must be less than the tolerance established (See Section 2-6).

The observation efficiency η_2 for the continuous drive method is directly affected by fluctuations in the mirror drive speed.

The signal-to-noise ratio S/N can be formulated by

$$S/N = \sqrt{(\eta_2 T)}\,[b(\sigma)\ \delta\sigma \pi \eta_1 A/R]/(NEP),$$

where the parameters are as follows:

η_2	observation efficiency
T	total time for the interferogram measurement between the zero path difference and the maximum path difference
$b(\sigma)$	radiative energy at σ
$\delta\sigma$	resolution
η_1	efficiency of the beam splitter
A	area of the interferometer mirror
R	resolving power
NEP	noise equivalent power of the detector

References

1. Penk, K., and Genzel, L. (1962) *Appl. Opt.* 1: 643.

2. Jacquinot, P., and Dufour, C., (1948) *J. Rech. du CNRS* **6**: 91.

3. Strong, J. (1958) *Concepts of Classical Optics*, W. H. Freeman, San Francisco.

4. Steel, W. H. (1964) *JOSA* **54**: 151.

5. Sakai, H., and Vanasse, G. (1966) *JOSA* **56**: 357.

6. James, D., and Ring, J. (1967) *J. de Physique* (*Supplement*) **28C2**: 150.

7. Loewenstein, E. V., and Englesrath, A. (1967) *J. de Physique* (*Supplement*) **28C2**: 153.

8. Roland, G. (1966), Thesis, Université de Liège.

9. Born, M., and Wolf, E. (1965), *Principles of Optics*, ed. 3 rev., Pergamon Press.

10. Mandel, L (1963) in *Progress in Optics*, Vol. 2, E. Wolf, Ed., North-Holland, Amsterdam.

11. Connes, J., and Connes, P. (1966) *JOSA* **56**: 896.

12. Connes, J., Connes, P., and Maillard, J. (1967) *J. de Physique* (*Supplement*) **28C2**: 120.

13. Mertz, L. (1965) *Transformations in Optics*, John Wiley.

14. Liu, B., and Kaneko, T. (1969) *Proc. IEEE* **57**: 1735.

15. Weinstein, C. J. (1969) *NEREM Record* 114.

16. Connes, J. (1961) *Rec. Opt.* **40**: 45, 116, 171, 231.

17. Hanel, R., et al (1969) *Appl. Opt.* **8**: 2059.

18. Sakai, H., and Murphy, R. (1970) *JOSA* **60**: 422.

19. Richards, P. (1964) *JOSA* **54**, 1474.

20. Connes, J., et al (1970) *Nouv. Rev. d'Opt. Appl.* 1: 3.

Bibliography

Fourier Spectroscopy

Connes, J. (1961) *Rev. Opt.* **40**: 45, 116, 171, 231.

Connes, P. (1970) *Am. Rev. Astro. Astrophys.*, in press.

Mertz, L. (1965) *Transformations in Optics*, John Wiley.

Vanasse, G., and Sakai, H. (1967), in *Progress in Optics*, Vol. 6, E. Wolf, Ed., North-Holland, Amsterdam.

Signal and Noise in the Measurement

Bracewell, R. M. (1965), *The Fourier Transform and Its Applications*, McGraw-Hill.

Jamieson, J. A., et al (1963), *Infrared Physics and Engineering*, McGraw-Hill.

Stratonovich, R. L. (1963) *Topics in the Theory of Random Noise* (Translated by R. Silverman) Gordon Breach, N.Y.

Multiplex Gain

Fellgett, P. (1951), Thesis, Cambridge University.

Vanasse, G., and Sakai, H. (1967), in *Progress in Optics*, Vol. 6, E. Wolf, Ed., North-Holland, Amsterdam.

Interferometric Optics

Steel, W. H. (1967), *Interferometry*, Cambridge Univ. Press, London.

Strong, J. (1958), *Concepts of Classical Optics*, W. H. Freeman, San Francisco.

Vanasse, G., and Sakai, H. (1967), in *Progress in Optics*, Vol. 6, E. Wolf, Ed., North-Holland, Amsterdam.

Mertz, L. (1965) *Transformations in Optics*, John Wiley.

Detector Noise

Jamieson, J. A., et al (1963), *Infrared Physics and Engineering*, McGraw-Hill.

Photon Noise

Jamieson, J. A., et al (1963), *Infrared Physics and Engineering*, McGraw-Hill.

Mandel, L. (1963) in *Progress in Optics*, Vol. 2, E. Wolf, Ed., North-Holland, Amsterdam.

Appendix 2-A.1

Suppose that the radiation field incoming to the interferometer has an amplitude U. All symbols are referred to those shown in Figure 2-A1. The field

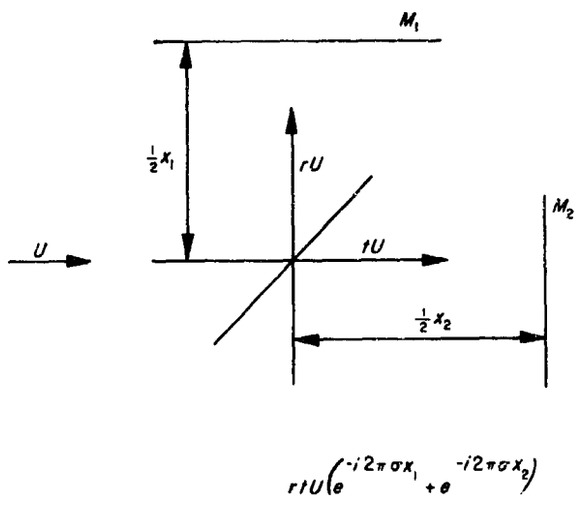


Figure 2-A1

going toward mirror M_1 has an amplitude of rU, where r is the amplitude reflection coefficient of the beam splitter. The field going toward M_2 is given by tU, where t is the amplitude transmission coefficient. The combined field is then given by

$$V = trU(e^{-i2\pi\sigma x_1} + e^{-i2\pi\sigma x_2}).$$

The terms inside the parentheses are the phase factors, the first for beam M_1 and the second for beam M_2.

The intensity is given by the product of V and its complex conjugate V^*:

$$VV^* = 2rr^*tt^*UU^*\{1+\cos 2\pi\sigma(x_1-x_2)\}.$$

The intensity coefficients of the reflection R and the transmission T are given by

$$R = rr^*;$$

$$T = tt^* = (1-R).$$

The efficiency is obviously given by

$$\eta = rr^*tt^* = RT = (1-R)R,$$

and its maximum value of 0.25 is found at $R=0.50$. Since the intensity of the incoming field is

$$UU^* = b(\sigma)\ \delta\sigma A\Omega,$$

the fringe intensity VV^* is given by

$$VV^* = 2\eta b(\sigma)\ \delta\sigma A\Omega\{1+\cos 2\pi\sigma x\},$$

where x is the path difference:

$$x = x_1 - x_2.$$

Appendix 2-A.2

The photon energy of the infrared radiation at 5000 cm^{-1} is given by

$$h\nu = hc\sigma = 10^{-12} \text{ erg or } 10^{-19} \text{ joule.}$$

The fluctuation in the photon arrival to the detector becomes larger than the detector noise of $NEP = 10^{-13} W \sec^{-\frac{1}{2}}$, if the number of arriving photons exceeds 10^{12}/sec. This figure is obtained in the following two steps of the calculation. First, the critical fluctuation $\sqrt{\langle \Delta n^2 \rangle}$ is calculated from

$$\sqrt{\langle \Delta n^2 \rangle} = NEP/h\nu = 10^{-13}/10^{-19} = 10^{6}.$$

Then, the photon intensity exhibiting this fluctuation is determined from

$$n = \{\sqrt{\langle \Delta n^2 \rangle}\}^2 = (10^6)^2 = 10^{12}.$$

The signal-to-noise ratio for this measurement,

$$S/N = n/\sqrt{\langle \Delta n^2 \rangle} = 10^6 \text{ per sec}^{\frac{1}{2}},$$

is unrealistically high for practical measurements. For this high a figure, there seems to be no way to determine the accuracy of the intensity measurement. If the signal-to-noise ratio is lowered to the practical accuracy of the measurement ($\sim 10^4$) by reducing the observation time, the entire interferogram must be measured during about 10^{-4} sec in time. This measurement seems impossible at present since the fringe frequency exceeds the cutoff of detectors now available.

Appendix 2-A.3

For the sinusoidal sampling errors Δx, assume

$$\Delta x = x' - x = \theta \sin 2\pi\varphi x.$$

Here x represents the true position and x' the actual sampling position. Obviously,

$$x' = x + \theta \sin 2\pi\varphi x.$$

The monochromatic interferogram $F(x')$ sampled with these errors is given by

$$F(x') = B \cos 2\pi\sigma_0 x'$$

$$= B \cos 2\pi(\sigma_0 x + \sigma_0 \theta \sin 2\pi\varphi x),$$

which can be expanded to

$$F(x') = B\{J_0(\sigma_0\theta)\cos 2\pi\sigma_0 x + J_1(\sigma_0\theta)[\cos 2\pi(\sigma_0-\varphi)x + \cos 2\pi(\sigma_0+\varphi)x] + \cdots\}.$$

If the argument $\sigma_0\theta$ in these Bessel functions is much smaller than 1, the Bessel functions can be expressed by

$$J_0(\sigma_0\theta) \approx 1,$$

$$J_1(\sigma_0\theta) \approx \tfrac{1}{4}\sigma_0\theta;$$

and all other terms of higher order are at least one order of magnitude smaller than the second term $J_1(\sigma_0\theta)$. Hence,

$$F(x') \approx B\{\cos 2\pi\sigma_0 x + \tfrac{1}{4}\sigma_0\theta[\cos 2\pi(\sigma_0-\varphi)x + \cos 2\pi(\sigma_0+\varphi)x]\},$$

showing that the Fourier transform of $F(x')$ consists of a central peak having a height B and two satellites of $\frac{1}{4}B\sigma_0\theta$ located at $\sigma_0 \pm \varphi$.

Appendix 2-A.4

The resolving power is given by

$$R = \sigma_0/\delta\sigma.$$

Thus, Eq. (2-40) of Section 2-6 can be rewritten as

$$\omega_\beta = \tfrac{1}{4}\omega_x\sigma_0(\sigma_0/\delta\sigma)^{\frac{1}{2}} = \tfrac{1}{4}\omega_x\sigma_0^{\frac{3}{2}}(\delta\sigma)^{-\frac{1}{2}}.$$

Setting the upper bound of ω_β to be 0.01, we write the corresponding sampling error ω_x, the tolerance, by

$$\omega_x = 0.04(\delta\sigma)^{\frac{1}{2}}\sigma_0^{-\frac{3}{2}} = 0.04/\sigma_0 R^{\frac{1}{2}}.$$

Appendix 2-A.5

The RC circuit shown in Figure 2-A2 has a filter characteristic given by

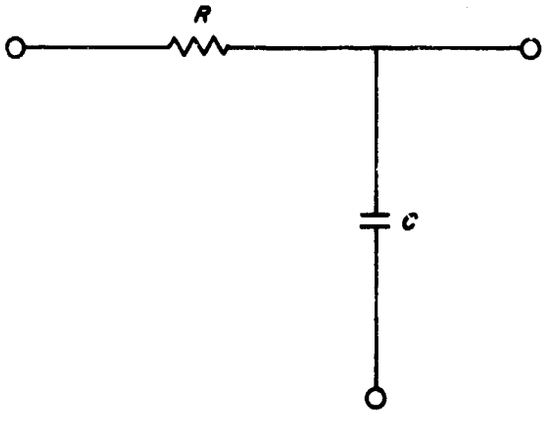


Figure 2-A2

$$\varphi(t) = \frac{1}{\tau_0} e^{-t/\tau_0},$$

where τ_0 is the time constant of the filter, given by

$$\tau_0 = 1/(RC) = \int_{-\infty}^{0} e^{-t/\tau_0}\, dt.$$

This equation indicates that the time constant represents the duration of effective integration. The trans-

fer function $\tilde{\varphi}(f)$ in the frequency domain, calculated from $\varphi(t)$, is:

$$\tilde{\varphi}(f)=\int_{-\infty}^{\infty}\varphi(t)e^{i2\pi ft}\,dt$$

$$=\frac{1}{\tau_0}\int_{-\infty}^{0}e^{-t/\tau_0}\,e^{i2\pi ft}\,dt$$

$$=\frac{1}{\tau_0}\,\frac{1}{i2\pi f-(1/\tau_0)}=\frac{1}{\tau_0}\left\{\frac{1}{4\pi^2f^2+(1/\tau_0)^2}\right\}^{\frac{1}{2}}$$

$$\times e^{i\tan^{-1}(2\pi f\tau_0)}.$$

With the transfer function expressed as

$$\tilde{\varphi}(f)=|\tilde{\varphi}(f)|e^{i\alpha(f)},$$

the gain can be written as

$$|\tilde{\varphi}(f)|=\frac{1}{\tau_0}\left\{\frac{1}{4\pi^2f^2+(1/\tau_0)^2}\right\}^{\frac{1}{2}}$$

and the phase as

$$\alpha(f)=\tan^{-1}(2\pi f\tau_0).$$

Values of the gain and phase for typical values of the frequency are listed in Table 2-A5.1.

Table 2-A5.1 Typical Values of Gain and Phase

frequency (f)	gain ($\tilde{\varphi}f$)	phase $\alpha(f)$	
		(radian)	(degree)
0	1.000	0.0	0.0
$1/2\pi\tau_0$	0.707	0.25π	45.0
$1/\pi\tau_0$	0.446	0.36π	64.5
$1/\tau_0$	0.158	0.45π	81.0

It is immediately apparent that owing to the phase α, this filter introduces a rather large asymmetry in the interferogram.

If the highest spectral frequency is σ_M in the signal bandwidth, the corresponding fringe frequency is given by

$$f_M=\sigma_M v,$$

where v is the speed of the interferometer mirror. Then the slowest sampling rate is $2f_M$. When we match the time constant τ_0 of the filter with this sampling rate $2f_M$, the highest frequency content in the spectrum is recovered with a gain factor of 0.158, a rather poor figure. A solution to this problem is found in an increase in the sampling rate from $2f_M$ to about $4\pi f_M$. The spectrum is then recovered without much attenuation, and the gain varies between 1 and 0.707 within the signal bandwidth.

Appendix 2-A.6

The tolerance ω_x derived in Appendix 2–A.4 is given by

$$\omega_x = 0.04/(\sigma_0\sqrt{R}),$$

where σ_0 is the optical frequency and R is the resolving power. Equating this expression for Δx with Eq. (2-44), we get

$$\tfrac{1}{2}\tau\,\Delta v = 0.04/(\sigma_0\sqrt{R}),$$

or,

$$\tau = 0.08/(\sigma_0\,\Delta v\sqrt{R}).$$

The total observation time is obtained by multiplying the total number of interferogram samplings by the time constant τ. Thus,

$$R\tau = 0.08\sqrt{R}/(\sigma_0\,\Delta v).$$

The total time available for the measurement is

$$T = X/v = 1/(2v\,\delta\sigma),$$

where X is the maximum path difference, related to the resolution $\delta\sigma$ by

$$X = 1/(2\,\delta\sigma).$$

Since the resolving power R is given by

$$R = \sigma_0/\delta\sigma,$$

the efficiency η is then given by

$$\eta = R\tau/T = 0.16\,\frac{v}{\Delta v}\,\frac{\delta\sigma}{\sigma_0}\sqrt{R} = \frac{0.16}{\sqrt{R}}\,\frac{v}{\Delta v}.$$

Contents

3. Interferometers for Fourier Spectroscopy

W. H. Steel
National Standards Laboratory
Sydney, Australia

Abstract

An interferometer for Fourier spectroscopy should have two interfering beams and an easy method of varying their path difference by moving some component. Accidental errors of this movement should not limit the source size, nor, preferably, should it be limited by image shifts accompanying the change of path. Polarization errors should be absent.

These criteria are discussed and applied to existing interferometers.

3–1 INTRODUCTION

The optical system of a Fourier spectrometer is a two-beam interferometer with a variable delay τ. A detector gives an output consisting of a uniform background signal on which is superimposed an oscillatory function of the delay, called the interferogram. This interferogram is the autocorrelation function of the radiation field at times τ apart, and hence the Fourier transform of the power spectrum of the radiation. The spectrometer may be used to measure the spectrum of an external source, or a refractive-index spectrum of a sample placed inside it.[1]

The simple theory[2] that leads to the exact Fourier transform considers a point source of radiation and a point detector of radiant power. Two rays connect these points, one following each arm of the interferometer. The transit time for radiation along each ray is given by the sum $\sum nd/c$, the ray length d in each medium through which it passes divided by the speed of light in that medium c/n. The difference between the times for the two paths is the delay τ of the interferometer. This multiplied by the speed of light is the optical path difference $x = c\tau$, a more appropriate quantity for a discussion of interferometers since lengths of arms are measured rather than transit times. The optical path difference is the difference of the sums $\sum nd$ for the two rays.

In an ideal symmetrical interferometer, the interferogram $F(x)$ is an even function of the path difference x and the spectrum is then given by a cosine transform

$$B(\sigma) = 2\int_0^\infty F(x)\cos 2\pi\sigma x\, dx; \qquad (3\text{-}1)$$

the reciprocal coordinate to path difference is the wave number σ. In practice the interferometer can

reach some finite maximum value X of the path difference and this determines the resolving power of the spectrometer; the smallest resolved interval is $\delta\sigma = 1/X$ and the resolving power is $R = \sigma/\delta\sigma = \sigma X$. The path difference should vary from zero to X, either continuously or in steps of a size chosen by the sampling theorem. If the interferometer produces an asymmetrical interferogram, correction of the phase error[3] requires that the path difference should also cover a small range of values on the other side of zero. This extra range gives no increase in resolving power.

The optical path difference is changed by increasing the distance travelled in some part of one arm, usually by movement of a mirror. Preferably the change is made in air or *in vacuo*, so that the optical path difference is independent of wave number. If the path is changed in a dispersive medium, the path difference varies with wave number and the spectrum is obtained on a nonlinear scale of wave number. This is a minor inconvenience that can be corrected during the computation of the transform.

The starting point in designing a suitable interferometer is to provide a path difference that can be varied conveniently over a range determined by the requirements of resolving power and phase correction. An initial choice of suitable types is made on the simple two-ray picture.

3-2 PRINCIPAL TYPES

Interferometers are classified traditionally into those with division of wavefront and those with division of amplitude. Both types are used in Fourier spectroscopy, but the latter has a much greater variety of forms.

3-2.1 Division of Wavefront

For division of wavefront, the two rays leave the source in different directions and arrive at the detector from different directions. In between, they are separated and the path of one may be varied easily by some means such as the movable mirror shown in Figure 3-1.

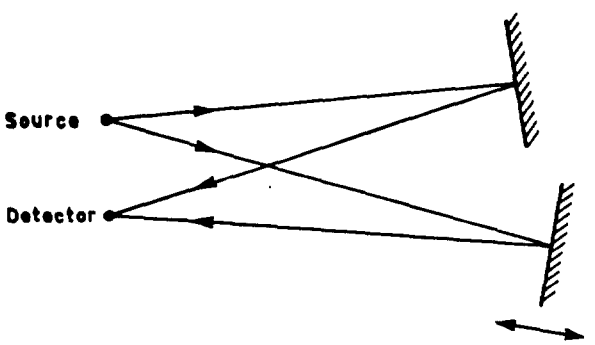


Figure 3-1. Simple Division-of-wavefront Interferometer with Variable Path Difference

3-2.2 Division of Amplitude

In division of amplitude, a single ray from the source is divided into two by a beam-splitter. These two proceed separately in a region where a variable path difference can be easily introduced, and then recombine at a second beam-splitter or at a second encounter with the first.

These interferometers can be further subdivided according to the type of beam-splitter that is used.

3-2.2.1 PARTIAL REFLECTORS

The most common beam-splitter is the partial reflector or half-silvered mirror that reflects part of the incident radiation and transmits part. At long wavelengths it is made from a stretched sheet of plastic film, usually Mylar[4]. For shorter wavelengths a thin coating of a dielectric (or occasionally a metal) on a transparent support is used.

The three basic two-beam interferometers with this type of beam-splitter are shown in Figure 3-2.

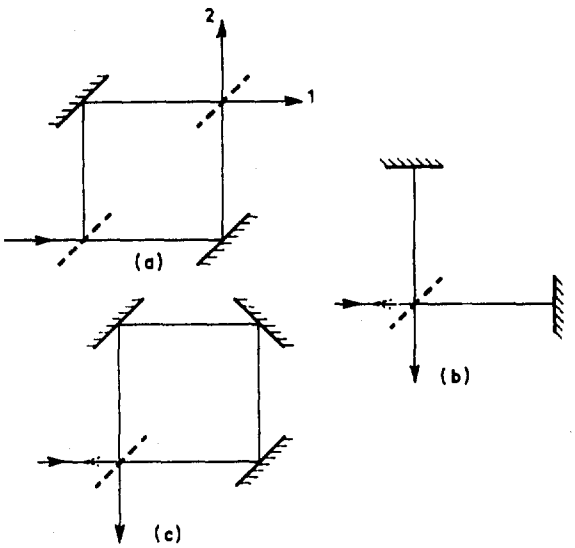


Figure 3-2. Basic Two-beam Interferometers with Division of Amplitude. (a) Mach-Zehnder, (b) Michelson (c) cyclic

The Michelson interferometer gives the simplest method of changing the path difference and it is the type most used in Fourier spectroscopy. The Mach-Zehnder interferometer has the advantage of two separate outputs in which the interferograms may be complementary, a feature that can be useful for reducing the noise from source fluctuations. It is not easy to change the path length, however, unless the modification shown in Figure 3-3(a) is used. In the Michelson interferometer the second output is returned to the source. It can be recovered if the

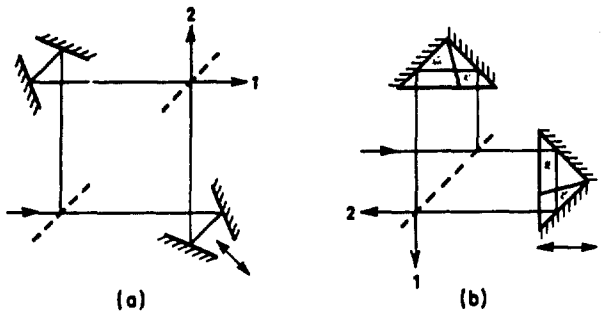


Figure 3-3. Modified Interferometers. (a) Mach-Zehnder with pairs of mirrors (b) Michelson with retroreflectors

modification (Figure 3-3(b)) is used, where the mirrors have been replaced by retroreflectors, which are shown as cube corners.

The cyclic interferometer is not a suitable form, since its path difference is not easily varied.

3-2.2.2 POLARIZATION INTERFEROMETERS

Partial reflectors are usually partial polarizers. Fully polarizing beam-splitters can be made of stacks of suitable dielectric layers, from birefringent materials or, for longer wavelengths, of wire grids.

A polarization interferometer, however, implies more than the use of a polarizing beam-splitter. It is an interferometer made of a birefringent material, such as quartz, calcite, or rutile, whose refractive index differs for the two orthogonal linear polarizations. Two rays may follow the same path yet have an optical path difference due to the index difference. The path difference can be varied by changing the effective thickness of the material.

3-2.3.3 DIFFRACTING BEAM-SPLITTERS

A diffraction grating can be regarded as a beam-splitter, since it divides an incident ray into one direct and several diffracted rays. It is difficult to make a grating that produces only two rays, but one that gives three rays—one direct and two diffracted—is feasible.

A Fourier spectrometer that uses a grating as beam-splitter is the mock interferometer of Mertz et al.[5]

3-3 FURTHER CONSIDERATIONS

Further factors that affect the choice of a suitable type of interferometer are the properties of beam-splitters and reflectors, and the action of the interferometer on polarized light.

3-3.1 Beam-splitters

A single-film or coating beam-splitter reflects radiation from both its surfaces. To give sufficient reflectance, these two reflected components should be in phase so that they reinforce; the film should be one quarter-wavelength thick.

Obviously, a beam-splitter can have optimum thickness for one wavelength only. For longer wavelengths its reflectance drops to zero as the wave number tends to zero; for shorter wavelengths it oscillates between zero and its maximum value as the thickness goes through even and odd multiples of $\lambda/4$. Around the optimum wavelength it has a useful range of two to three octaves; for larger spectral ranges the beam-splitter would need to be changed.

A beam-splitter that is a coating on a transparent support requires a *compensating plate* in the other beam, of the same material and thickness as this support. This enables the optical paths in the two beams to be matched at all wavelengths. Without such a match, an unsymmetrical interferogram is obtained, which requires phase correction to give the spectrum.

Partially reflecting beam-splitters are partial polarizers, reflecting more of one linear polarization than the other. This in itself is not serious because a beam-splitter is usually used twice, once in transmission and once in reflection, and the product of reflectance and transmittance varies much less than each factor alone. But the polarizing effects of a beam-splitter can interact with those of other components of the interferometer.

3-3.2 Polarization Effects

A polarization interferometer is used with a polarizer on each side of it. It thus analyses one linear polarization of the source at a time.

Interferometers with partial reflectors, however, analyse both polarizations together. The two interferograms, produced by the two principal directions of linear polarization, should coincide or modulation is lost. Since oblique mirrors give a phase difference to these two polarizations, their use can reduce modulation.

If similar oblique mirrors are used in the other arm of the interferometer, these phase differences cancel provided all components, mirrors and beam-splitter, have the same principal directions. But if the mirrors are skewed, this cancellation does not occur,[6] since the result of a sequence of polarizing components at different azimuths depends on their order and this cannot be the same for both beams (in one beam the beam-splitter reflects first, in the other it transmits first).

3-3.3 Reflectors

Two types of reflector are used in interferometers: specular reflectors and retroreflectors.

A plane mirror is a specular reflector and the reflected ray leaves from the point of arrival of the incident ray at an equal angle but on the opposite side of the normal. In a retroreflector, the reflected ray leaves at the same angle as the incident ray but from a point an equal distance on the opposite side of the center to the point of incidence. The reflected ray is unaffected by movements of a specular reflector in

its own plane but changes angle if the reflector is tilted. The ray from a retroreflector is unaffected by tilts but is displaced by lateral movements of the reflector.

Figure 3-4 shows three forms of retroreflector. The

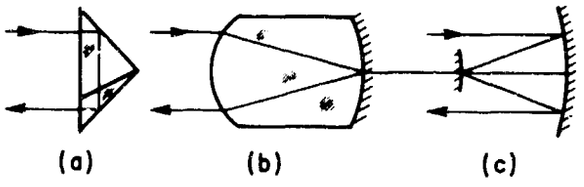


Figure 3-4 Retroreflectors. (a) cube corner, (b) refracting cat's eye (c) all-reflecting cat's eye

first, the cube corner, is made either as a prism or as three mutually perpendicular mirrors. The others are known as 'cat's eyes', both refracting and reflecting.

The mirrors of a cube corner are skewed, and an interferometer using these will show polarization effects whose magnitude depends on the optical constants of the mirror and beam-splitter coatings. Cube corners are successfully used in Fourier spectrometry, so the loss of modulation of the interferogram is not always serious, but it is a factor to be watched. The mirrors of a cat's eye are used at close to normal incidence and polarization effects are very small.

The large, concave mirror of the cat's eye is commonly made a paraboloid to correct for spherical aberration; the coma of a single paraboloid is corrected when the same mirror is used twice. However, spherical mirrors are simpler to produce and the spherical aberration is the same for both beams and so cancels. A second-order error occurs when the cat's eye is tilted[7], but this is usually smaller than the source-size errors discussed next.

3-4 FINITE BEAMS

3-4.1 Beam Size

A single ray transmits no radiant power. Finite power requires a beam of finite area of cross section and finite solid angle of spread. The power that such a beam conveys is directly proportional to the product of area and solid angle, the *étendue* or optical extent of the beam. This product is invariant along the beam as it passes through an optical system.

In an image-forming system, the *étendue* of the beam is fixed by two apertures, the entrance pupil and the field stop. There are similarly two limiting apertures in an interferometer; these may be the source, the detector, the beam-splitter, or the mirrors. In interferometers in which fringe patterns are viewed, such as those for testing lenses, one aperture determines the size of the interference pattern, the field aperture; the other is the aperture of the effective source. In Fourier spectrometers there is no such distinction between the two apertures; all that matters is the total radiation between them. However, it is convenient to retain the names, although one limiting aperture may be, for example, the detector rather than the source.

3-4.2 Variations of Path Difference

For the same theory to apply for two extended beams as for single rays, the optical path differences for a pair of rays from any point in the source plane to any point in the field plane should be the same; for the final interferogram is derived from the integrated interferograms for all such pairs of rays and a variation of path difference will reduce the modulation of this sum.

Figure 3-5 shows two rays through the interferom-

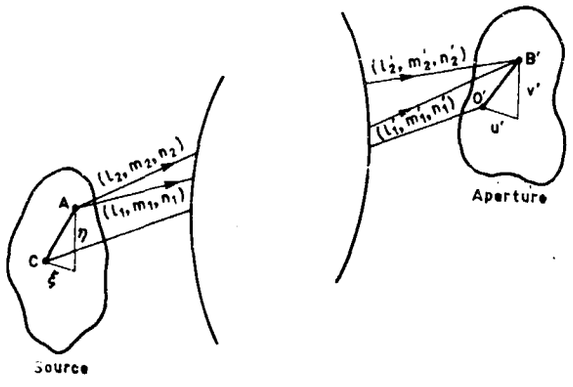


Figure 3-5 Positions at Reference Planes and Directions of Rays

eter that connect a point (ξ, η) at the source to a point (u', v') in the field plane. These rays leave the source with direction cosines (l_1, m_1, n_1) and (l_2, m_2, n_2) and arrive at the field plane with direction cosines (l_1', m_1', n_1') and (l_2', m_2', n_2').

For neighboring points in the two planes, the variation in path is given by Fermat's principle as

$$dP = l'\,du' + m'\,dv' - l\,d\xi - m\,d\eta \qquad (3\text{-}2)$$

for both beams. The variation in path difference is thus

$$dx = (l_2' - l_1')\,du' + (m_2' - m_1')\,dv' - (l_2 - l_1)\,d\xi - (m_2 - m_1)\,d\eta. \qquad (3\text{-}3)$$

For the path difference to be the same for all rays, its variation dx must be zero for any variations du', dv', $d\xi$, or $d\eta$. Thus every pair of rays must arrive at the field plane from the same direction and leave the source from the same direction. If this is not so, the corresponding aperture, field or source, must be made sufficiently small so that the finite variation δx is tolerable, which means in practice $\delta x < \lambda$, where $\lambda = 1/\sigma$ is the wavelength of the radiation.

3-4.2.1 INTERFEROMETER PARAMETERS

To separate the various effects that cause the path variations, it is convenient to introduce four parameters[6] to describe the interferometer. Figure 3-6

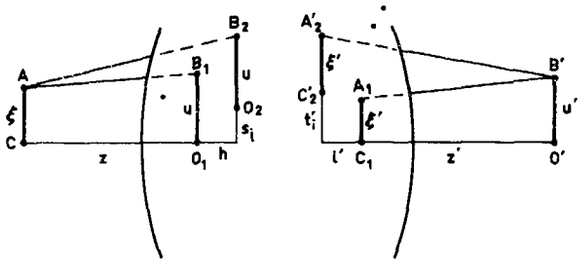


Figure 3-6. One Section of Figure 3-5, Showing Images of Each Plane as Seen From the Other

shows, in two dimensions only, an elaboration of Figure 3-5 to which have been added the images of each plane as seen from the other. Each image is doubled, there being one image for each beam, and these two images may be separated laterally and longitudinally. Either separation can cause a difference in direction of a pair of rays.

For the field images, the lateral separation is the shear s, and the longitudinal separation is the shift h. For the source images, these are the tilt t' and the lead l'. The difference between direction cosines can be expressed in terms of these parameters and of the two separations z and z' from source to aperture image and from source image to aperture, to give the variation of path difference dx. Integration gives the path difference, which is found to be

$$x = x_0 - (t_i'u' + t_j'v')/z' - l'(u'^2+v'^2)/2z'^2 - (s_i\xi + s_j\eta)/z - h(\xi^2+\eta^2)/2z^2, \qquad (3\text{-}4)$$

where t_i', t_j' and s_i, s_j are the two normal components of the tilt and shear and x_0 is the path difference for a central pair of rays.

If a shear is present, the variation $x - x_0$ is small only when the source is small in the direction of the shear (a narrow slit); if a shift is present the source should be small in all directions (a small circle). A tilt or a lead imposes similar conditions on the field size.

3-4.2.2 CHOICE OF APERTURE SIZE

Although not designed for the purpose, a Fourier spectrometer can show interference fringes when monochromatic radiation is used with it, provided the beam-splitter transmits some visible radiation. Interference patterns can be observed in both the source and field planes in turn when the aperture in the other plane is made sufficiently small. The fringes for a Michelson interferometer are shown in Figure 3-7.

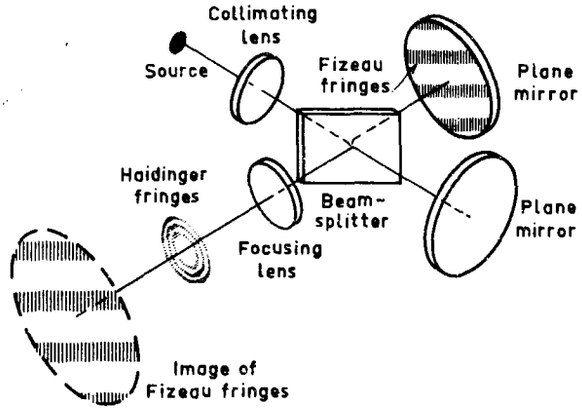


Figure 3-7. Two Sets of Fringes in a Michelson Interferometer

Interference fringes are contours of each one-wavelength change of path difference. If each aperture of the interferometer is made to cover no more than one half a fringe spacing of the fringes that can be seen in its plane, rays between two such apertures cannot have a variation of path difference exceeding one wavelength.

3-5 COMPENSATION

It is possible to make more complicated interferometers in which one or more of the four parameters is always zero. This technique is called interferometer compensation.

3-5.1 Sources of Errors

There is no reason, in theory, why a Fourier spectrometer should have either a shear or a tilt. In practice, however, these occur because of the difficulty of moving a mirror (or similar component) to change the optical path without accidentally tilting it or displacing it laterally. The correction of shears and

tilts is thus a balance between the merits of compensation vs better mechanical design and construction. It is usually possible to have only one of the two present. if the mirror moved is at an image of the aperture plane, this image is not affected by a tilt of the mirror.

The problem of shears and leads is different. As they are longitudinal effects, a change of optical path in a simple interferometer will ordinarily produce a change of these two parameters, and this can be avoided only by a compensation technique. Again it is possible to eliminate one but not both effects by simple means. In the common form of Michelson interferometer there is a collimating system so that, from the aperture plane, the source appears at infinity. Any finite change of distance, as used to change the path difference, will not be noticed; this interferometer will have a shift but no lead. This means that the size of the source only is limited. The aperture, as given by the size of the beam-splitter or mirrors, can be as large as space or money allows.

3–5.2 Scanning Functions for No Compensation

The integrated effect of variations in path difference across the beam is to reduce the modulation of the interferogram. This reduction sets the maximum values of the avoidable parameters, tilt and shear, that can be allowed.

In an uncompensated interferometer, there is an unavoidable variation of path difference, expressible in terms of the longitudinal parameters, that increases with the path difference x. This leads to a modulation that can be expressed as the product of an ideal interferogram and a decreasing function of x, sometimes called an aperture function. The resulting spectrum is degraded by convolution with a scanning function that is the Fourier transform of this aperture function.

Thus, phase variations that increase with path difference degrade the spectrum in the same way that aberrations in a classical spectrometer do. Aberrations of an interferometer, and other constant phase variations, do not cause any such degrading, although they do reduce the signal obtained.

Two aperture functions have been calculated; both are complex functions. Their absolute values are shown in Figure 3-8.

The first case is that of an interferometer in which one limiting aperture, source or detector, is at infinity. If the source is circular and subtends a solid angle Ω, the aperture function is

$$\operatorname{sinc}\left(\frac{\Omega\sigma x}{2\pi}\right)\exp\left(-\frac{1}{2}i\Omega\sigma x\right);\ \operatorname{sinc} x=\sin\pi x/\pi x. \quad (3\text{-}5)$$

The interferogram thus disappears at the first zero of this function, $x=2\pi\ \Omega\sigma$. Hence the maximum path difference X must be less than this value: the source size is limited by

$$\Omega<2\pi/\sigma X=2\pi/R, \quad (3\text{-}6)$$

where R is the resolving power.

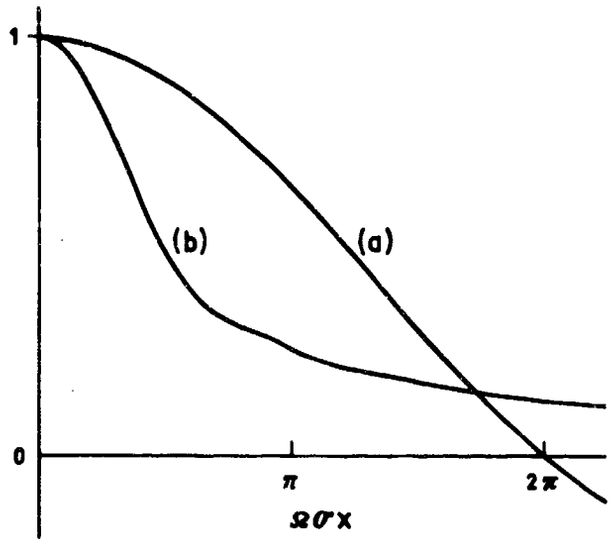


Figure 3-8. Absolute Values of Aperture Functions for an Interferometer: (a) With Collimation and a Circular Source (b) With no Collimation and Two Circular Apertures of the Same Size[8]

The second case given is for an interferometer without collimation. Both apertures are circles of the same size. The modulation falls more rapidly than for an interferometer with collimation, but never reaches zero.

3–5.3 Tilt and Shear Compensation

If a moving mirror of a Michelson interferometer is accidentally tilted, it introduces a tilt and the modulation of the interferogram is reduced. If a retroreflector is used instead,[9,10] tilts have no effect but accidental lateral displacements produce a shear, which has a similar effect on the interferogram. In practice, it is simpler to meet the tolerance on lateral displacements than that on tilts, and retroreflectors have a practical advantage over plane mirrors. A combination of a moving retroreflector and a stationary mirror, shown in Figure 3-9, is insensitive to both tilts and displacements of the moving component and so is fully compensated.

Another method of compensation is to move together two components, one in each arm of the interferometer, so that one path is increased as the other is reduced. These components can be coupled so that errors of movement cancel. Two mirrors are

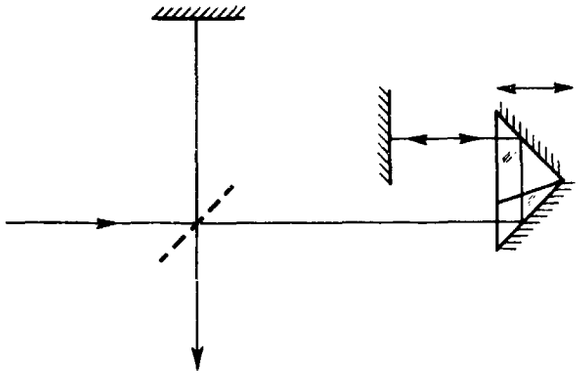


Figure 3-9. A Combination of Movable Retroreflector and Stationary mirror that Introduces Neither Shear nor Tilt

most simply coupled by placing them back to back. For the tilts in both arms to be the same as this mirror pair is ⁻ted, the beams in each arm must twist, so that, joined together, they have the topology of a Möbius band with 180° twist. An example of such an interferometer is shown in Figure 3-10, it can be

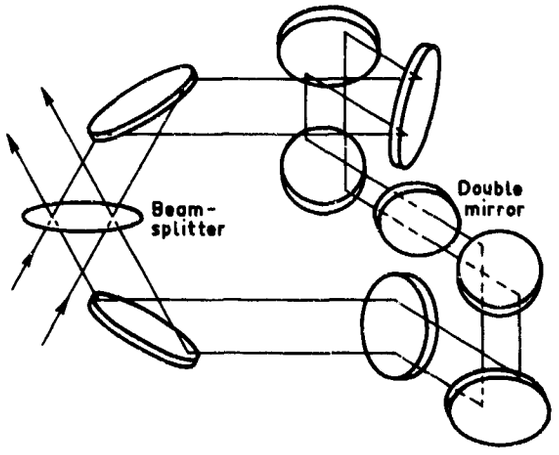


Figure 3-10. A Möbius-band Interferometer Corrected for Polarization Effects

made with fewer mirrors but at the expense of introducing polarization errors.

Similar compensation is possible with two retroreflectors joined together. Now the beams should have no twist: a simple 'all-in-one-plane' interferometer is compensated.[14]

3-5.4 Field-Widening

The compensation for shifts and leads is more complex. In a simple interferometer these parameters are directly related to the path difference, yet we require a variable path difference but no shift nor lead. This is possible with certain special designs of interferometer, said to be field-widened since the source size is no longer limited to $\Omega < 2\pi/R$.

Two main methods are used to field-widen an interferometer. The first involves changing the length of path in two different media simultaneously, for example air (or vacuum) and glass (or a transparent crystal), as shown in Figure 3-11. Compensation is possible since the optical path in a medium of thickness d and refractive index n is nd while an image seen through it appears at a distance d/n.

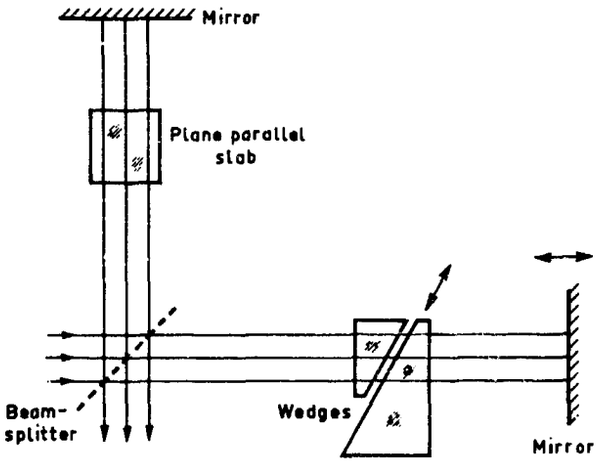


Figure 3-11. By Suitably Coupled Movements of the Wedges and the Mirror the Optical Path Can be Varied Without Introducing a Shift

The second method involves changing a focusing system inside the interferometer. Connes[15] has shown that if a telescope is placed in each arm of a Michelson interferometer, a movement of one telescope and one mirror together can give field-widening. For Fourier spectrometers, a more attractive method is that suggested by Cuisenier,[16] Cuisenier and Pinard,[7] and Erickson[14] of changing the curvature of a mirror, such as the convex mirror of a cat's eye, as the optical path is being changed.

Field-widening does not allow unlimited source sizes, since second-order effects (aberrations of the components) introduce their own limit. However, gains in solid angle of about 100 times that for a collimated interferometer are possible.

3-6 CATALOG OF INTERFEROMETERS

It is now appropriate to list the main interferometers used for Fourier spectroscopy to see how they meet the different requirements that have been discussed. Most systems are variants of the Michelson interferometer, but there are first three special types.

3-6.1 Lamellar Grating

The lamellar grating, Figure 3-12, consists of a

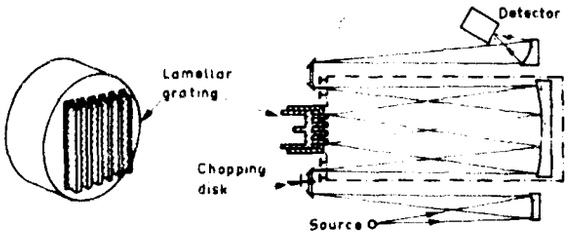


Figure 3-12. Lamellar Grating,[18,17]

fixed, slotted mirror and through these slots the tongues of the movable mirror can pass. It is a multiplication of the simple interferometer of Figure 3-1 and is the sole example of an interferometer with division of wavefront.

As the rays must leave the source and arrive at the detector at different angles, a shear is always present and the source must be a narrow slit. This means that the *étendue* is less than that possible with division of amplitude.

The lamellar grating is an instrument for the far infrared, since it is difficult to make mirrors of sufficient accuracy for shorter wavelengths. The reduced *étendue* is then not serious, since detector size is often the limitation in this region; larger beams could not be used. Richards[19] has shown that it is more efficient than a Michelson interferometer for wavelengths greater than 0.1 mm ($\sigma < 100\ \mathrm{cm}^{-1}$) and it has a larger useful range, since there are no beam-splitters that may need changing.

3-6.2 Special Beam-Splitters

3-6.2.1 POLARIZATION INTERFEROMETERS

Alternative forms of a polarization interferometer are shown in Figure 3-13. The thickness of a birefringent plate is changed by making it as two wedges that can be slid past each other. A bias plate or a second pair of wedges with their optic axis at right angles to that of the first enables the path difference to reach zero.

Polarization interferometers are limited to spectral regions in which transparent birefringent crystals are available. The change of path is less (for quartz, much less) than the movement of the wedges. This makes the interferometer stable (relatively unaffected by tilts, and so forth) but of low resolving power.

In use, there must be a polarizer on each side of the birefringent plates. Sinton[21] has shown how, by replacing these by polarizing beam-splitters, the two polarizations are analysed simultaneously and no radiation is wasted.

The allowed source size for a polarization interferometer is less than that for an uncompensated Michelson interferometer by the factor

$$\frac{n^2}{\sqrt{2}}\,\frac{d}{D+d}, \tag{3-7}$$

where n is the mean refractive index of the birefringent material, d is the change of thickness required to produce the maximum path difference X, and D is the total thickness of the wedges and bias plate at zero path difference.

3-6.2.2 MOCK INTERFEROMETER

The mock interferometer (Figure 3-14) consists

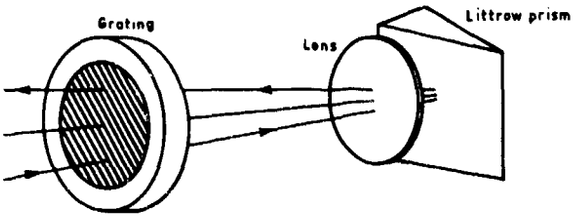


Figure 3-14. The Mock Interferometer of Mertz *et al.*[5] It has been developed further by Selby[22] and Ring and Selby[23]

of a Littrow monochromator in which the entrance and exit slits have been replaced by a rotatable grating. As this rotates, the diffracted beams rotate around the direct beam and, except for the wave number σ_0 which focuses back to the center of rotation, the path of the diffracted rays changes. Provided the spectrum studied lies wholly on one side of σ_0, an interferogram is produced that can be transformed to a spectrum. This has a nonlinear scale of σ with a displaced zero at σ_0.

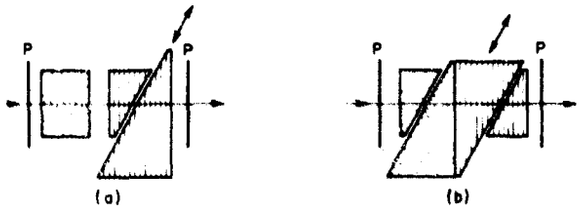


Figure 3-13. Polarization Interferometers: (a) Fixed Bias Plate[20] (b) Double Wedges

3-6.3 Michelson Interferometer

3-6.3.1 UNCOMPENSATED INTERFEROMETER

The simple Michelson interferometer, in spite of its sensitivity to mirror tilt, has the considerable advantages of compactness and simplicity. It is widely used in the far infrared[24] where, because tolerances are related to the wavelength, it is easier to make the mirror movement sufficiently accurate not to be troubled by tilts. However, it can be used satisfactorily for the near infrared, and Mertz[25] has made it in a particularly compact and stable form.

The aperture functions shown in Figure 3-8 indicate the advantage of using this interferometer (and others listed later) with collimating optics so that the aperture limiting the source, at either the source itself or at the detector, appears at infinity.

3-6.3.2 TILT-COMPENSATED INTERFEROMETER

The tilt-compensated system of Figure 3-9 is a modified Michelson interferometer in which both arms still end with plane mirrors. It has been used in the near infrared (3-4μm) by Rundle[26] with a cube corner as retroreflector.

The Möbius-band interferometer of Figure 3-10 has also been made for the near infrared by Pritchard *et al.*[13] It is quite insensitive to tilts of the moving mirror pair but has unfortunately proven to be much more sensitive to temperature changes (which tilt the fixed mirrors) than a simpler interferometer. Thus, a departure from a symmetrical construction has hidden disadvantages.

A different type of tilt-compensated interferometer, that remains a simple, plane instrument, is shown in Figure 3-15. Here the beam-splitter and a mirror rotate together as a parallel pair that compensates for tilts, which are here introduced deliberately. This rotation introduces the path difference.

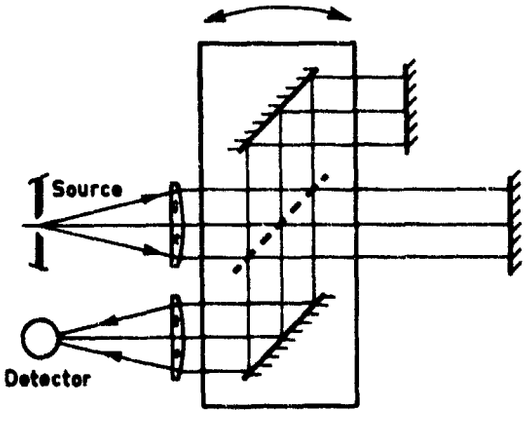


Figure 3-15. Tilt-compensated Interferometer of Sternberg and James[27]

3-6.3.3 FIELD WIDENED

A simpler form of the field-widened Michelson than that of Figure 3-11 is shown in Figure 3-16.

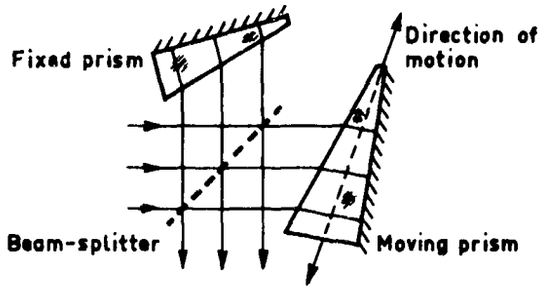


Figure 3-16. Field-widened Interferometer of Bouchareine and Connes (1963)[28]

Each wedge is now in a separate arm of the interferometer in the form of a prism with a reflecting coating on its back. One prism is moved in the direction of the image of this coating as seen from the front of the prism; this image then remains at a fixed distance as the optical path changes. The astigmatism of the prisms limits the gain in solid angle to 2 $\cot^2 \theta$, where θ is the angle of the prisms.

3-6.4 Michelson With Retroreflectors

If the plane mirrors of a Michelson interferometer are replaced by retroreflectors, the mechanical problem of changing the path without introducing errors is simplified. The retroreflectors can be used centrally in the beams, as are the mirrors in a Michelson, or the beams can enter on one side and leave on the other. This latter method requires a retroreflector having more than twice the beam diameter but it has the advantage of allowing two complementary outputs to be recorded. The extra parts of the aperture of the retroreflector can be used for additional beams, such as a monochromatic beam to monitor the path difference (measure it in terms of the wavelength of this radiation) and a beam of white light to indicate the position of zero path difference.

3-6.4.1 CUBE CORNERS

The retroreflectors used most commonly, for example by Fellgett,[29] and Roland[30] are cube corners, made of three mutually perpendicular mirrors. These are simple and fairly compact, but have the disadvantage of having polarization effects, although these do not appear to be serious enough to prevent fairly widespread use of this system. They are also rather delicate to adjust, for the mirrors should be accurately perpendicular to each other, particularly when the cube corner is used centrally in the beam.

3.6.4.2 CAT'S EYES

To avoid the polarization effects of cube corners and the delicate adjustment required, Connes[31] has used cat's eyes as retroreflectors. These are somewhat more bulky but have proven very satisfactory in his three high-resolving-power interferometers. They consist of a concave paraboloidal mirror with a convex sphere at the focus, the radius of the sphere being such as to image the beam-splitter back on itself (usually at zero path difference).

The principle of the interferometer is shown in Figure 3-17. The interferometer has been designed

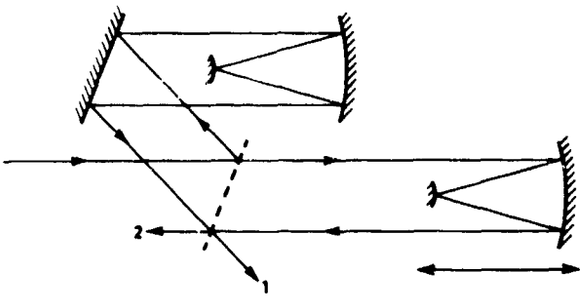


Figure 3-17. Interferometer With Cat's Eyes as Retroreflectors,[31,32]

for a servo-controlled step-by-step change of optical path. As the cat's eyes are massive, the first model displaced them continuously and produced the stepping by giving a saw-tooth displacement to the extra mirror. Connes' latest model,[32] however, has the cat's eyes stepped directly. It represents the summit of Fourier spectrometers for the near infrared, designed to give a resolving power of one million for one million spectral elements and at the same time, to have a very accurate scale of wave number. To reduce errors caused by source fluctuations and atmospheric scintillation, the optical path is modulated, as suggested by Mertz,[25] so that the interferogram is a sine rather than a cosine transform of the spectrum. This modulation is produced by vibrating the small mirror of one cat's eye.

3.6.4.3 COMPENSATED FORMS

If the two retroreflectors are moved together, as shown in Figure 3-18, errors due to accidental lateral displacements are compensated at the expense of having a heavier moving part in the interferometer. Alternatively, the two retroreflectors may be connected back-to-back, as in the two interferometers suggested by Erickson,[14] shown in Figure 3-19.

Interferometers with cat's eyes can be field-widened by the method suggested by Erickson[14] and Cuisenier[15] of deforming the convex mirror of one or both

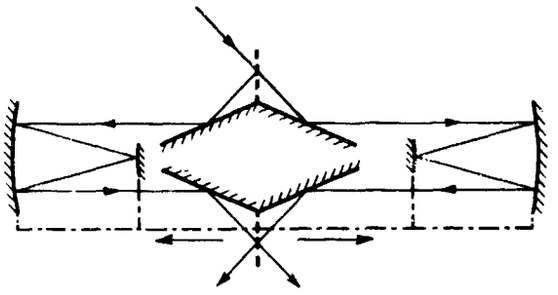


Figure 3-18. Method of Compensating a Cat's-eye Interferometer for Shear

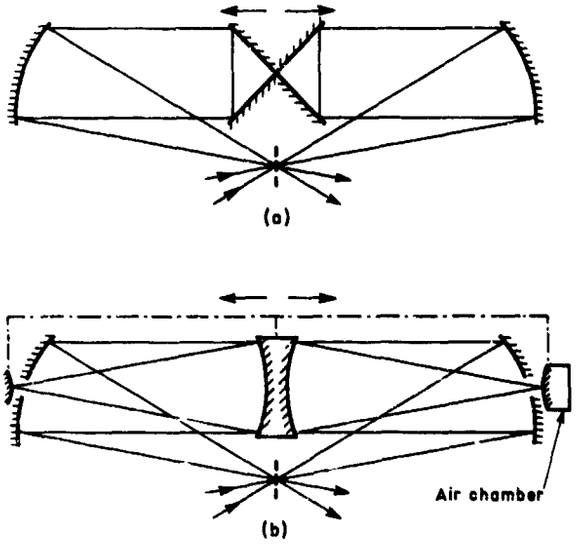


Figure 3-19. Compensated Interferometers With Two Retroreflectors Back to Back. a) two cube corners, b) two cat's eyes, one with a deformable mirror. Both interferometers have a small beam-splitter in convergent radiation. (Reproduced by permission of Keuffel and Esser Company)

arms. Although this has not yet been incorporated by Connes,[32] he has designed an instrument with the possibility of adding this refinement later, if needed. The distortion required is small; if the edge of the mirror is moved N wavelengths with respect to its center, the solid angle can be increased by a factor $2N$. In the interferometer in Figure 3-19(b) it was intended to change the curvature of one of the small mirrors by reducing the pressure in an air chamber behind it.

3-7 CONCLUSIONS

The final interferometer design chosen will depend very much on the use to which the instrument will be put. Important considerations are the spectral region to be studied, the resolving power required, and the

radiance of the source. The design will represent a compromise between simple systems with high mechanical accuracy and more complex interferometers in which mechanical errors are compensated, between the two extremes of the simple Michelson interferometer and the two moving cat's eyes of Figures 3-18 and 3-19(b) field-widened by deformation of the convex mirrors.

The last designs have not yet been made, but these and other complex systems may be tried in the future. Another possible development is the correction of mirror tilts by servo control rather than by mechanical precision or interferometer compensation. Instead of one auxiliary beam of monochromatic radiation to monitor path difference, three would be used to detect tilts as well.

References

1. Chamberlain, J. E., Gibbs, J. E., and Gebbie, H. A. (1963) *Nature* **198**: 874.

2. Vanasse, G. A., and Sakai, H. (1967) *Progress in Optics*, E. Wolf, Ed., **6**: 259. North Holland, Amsterdam.

3. Sakai, H., Vanasse, G. A., and Forman, M. L. (1968) *J. Opt. Soc. Am.* **58**: 84.

4. Gebbie, H. A. (1961) *Advances in Quantum Electronics*, J. R. Singer, Ed., Columbia Univ. Press, New York. p. 155.

5. Mertz, L., Young, N. O., and Armitage, J. (1962) *Optical Instruments and Techniques*, K. J. Habell, Ed. Chapman and Hall, London. p. 51.

6. Steel, W. H. (1967) *Interferometry*, Cambridge University Press.

7. Cuisenier, M., and Pinard, J. (1967) *J. Phys. (Paris)*, Suppl. **28**: C2-97; *published also as Méthodes Nouvelles de Spectroscopie Instrumentale*, C.N.R.S., Paris.

8. Steel, W. H. (1964) *J. Opt. Soc. Am.* **54**: 151.

9. Peck, E. R. (1948) *J. Opt. Soc. Am.* **38**: 66, 1015.

10. Peck, E. R. (1957) *J. Opt. Soc. Am.* **47**: 250.

11. Terrien, J. (1959) *Revue Opt. Theor. Instrum.* **38**: 29.

12. Murty, M.V.R.K. (1960) *J. Opt. Soc. Am.* **50**: 83.

13. Pritchard, J. L., Sakai, H., Steel, W. H., and Vanasse, G. A. (1967) *J. Phys. (Paris)*, Suppl. **28**: C2-91; *published also as Méthodes Nouvelles de Spectroscopie Instrumentale*, C.N.R.S., Paris.

14. Erickson, K. E. (1965) Private communication.

15. Connes, P. (1956) *Revue Opt. Théor. Instrum.* **35**: 37.

16. Cuisenier, M. (1964) *Etude et réalisation d'un intérferomètre de Michelson permettant l'utilisation de la spectroscopie par transformation de Fourier à haute resolution sur un intervalle spectral étendu.* Thèse de 3e cycle, Paris.

17. Strong, J. (1954) *J. Opt. Soc. Am.* **44**: 352.

18. Strong, J., and Vanasse, G. A. (1960) *J. Opt. Soc. Am.* **50**: 113.

19. Richards, P. L. (1964) *J. Opt. Soc. Am.* **54**: 1474.

20. Mertz, L. (1958) *J. Phys. Radium (Paris)* **19**: 233.

21. Sinton, W. M. (1963) *J. Quant. Spectrosc. Radiat. Transfer* **3**: 551.

22. Selby, M. J. (1966) *Infrared Phys.* **6**: 21.

23. Ring, J., and Selby, M. J. (1966) *Infrared Phys.* **6**: 33.

24. Gebbie, H. A., and Twiss, R. Q. (1966) *Rep. Prog. Phys.* **29**: 729.

25. Mertz, L. (1965) *Transformations in Optics*, Wiley, New York.

26. Rundle, H. N. (1965) *J. Res. Natl. Bur. Stand.* (U.S.) **69C**: 5.

27. Sternberg, R. S., and James, J. F. (1964) *J. Sci. Instrum.* **41**: 514.

28. Bouchareine, P., and Connes, P. (1963) *J. Phys. Radium* **24**: 134.

29. Fellgett, P. (1958) *J. Phys. Radium (Paris)* **19**: 237.

30. Roland, G. (1955) Spectroscopie par transformation de Fourier, Thèse, Liège.

31. Connes, J., and Connes, P. (1966) *J. Opt. Soc. Am.* **56**: 896.

32. Connes, P. (1969) *ICO Conf. Optical Instruments and Techniques*, (to be published) by Oriel Press, Newcastle-upon-Tyne.

Contents

4. Double-Beaming in Fourier Spectroscopy

Jerome M. Dowling
Space Physics Laboratory
Aerospace Corporation
El Segundo, California

Abstract

In two-beam interferometry, it is not the spectrum which is measured, but rather a quantity called the interferogram is obtained. The interferogram is related to the spectrum by a Fourier transformation. Because of this difference in what is measured, there are a number of problems encountered in Fourier spectroscopy which are of no consequence in conventional single slit dispersive spectroscopy. One of these is the dynamic range problem, which can be very serious in absorption type Fourier spectroscopy. Another is the unfamiliar form the raw data (that is, the interferogram) take. These and other problem areas are illustrated and remedial procedures are discussed.

4-1 INTRODUCTION

The two principal advantages of Fourier spectroscopy over conventional single slit scanning spectroscopy are a larger throughput and the multiplex advantage. However, there are a number of problems in Fourier spectroscopy which are not encountered in conventional single slit dispersive spectroscopy. One of these is the dynamic range problem, which can be very serious in absorption type Fourier spectroscopy. Another is the peculiar form the raw data take. It is not the spectrum which is measured, but rather a quantity called the interferogram, which is related to the spectrum by a Fourier transformation. It is the purpose of this paper to examine these and other problems and illustrate how they may be overcome or circumnavigated.

There are two types of interferometers that are in common use in the infrared and these are illustrated in Figure 4-1. Shown on the left is the conventional Michelson two-beam interferometer, which utilizes division of amplitude to obtain the interfering beams. Shown on the right is the lamellar grating interferometer which utilizes division of wavefront to obtain the two interfering beams. In either case, the signal produced by the detector as a function of optical path difference may be written[1]

$$D(x) = \int_{\sigma_1}^{\sigma_2} B(\sigma)(1 + \cos 2\pi\sigma x)\, d\sigma. \qquad (4\text{-}1)$$

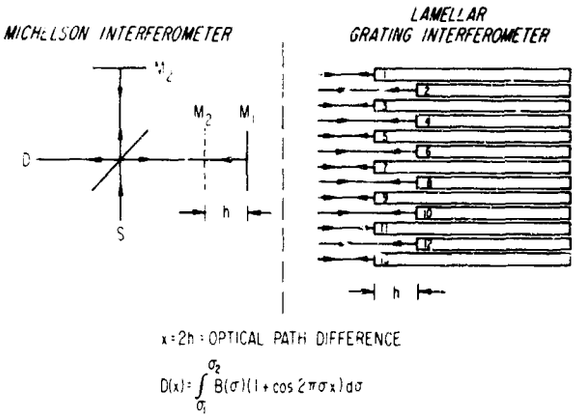


Figure 4-1. The Michelson and Lamellar Grating Interferometers

where $B(\sigma)$ is the spectral distribution of polychromatic radiation extending from σ_1 to σ_2 cm^{-1}, and x is the optical path difference. The mathematical operations required to recover $B(\sigma)$ from $D(x)$ involve a Fourier transformation which is discussed in detail in Chapters 6 and 8.

4–2 THE DYNAMIC RANGE PROBLEM

Shown in Figure 4-2 is a simple model of an absorption spectrum. The spectrum extends from σ_1 to σ_2 cm^{-1} and consists of triangularly shaped lines equally spaced in wavenumber units. The absorption lines remove approximately 4 percent of the radiation between σ_1 and σ_2.

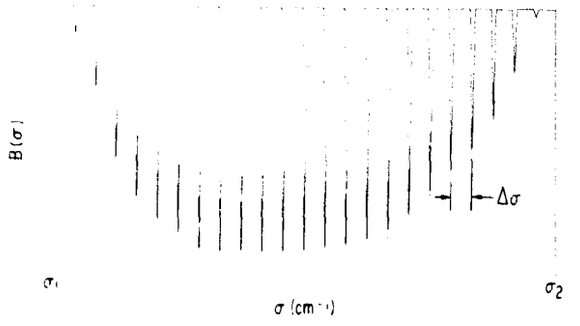


Figure 4-2. A Spectrum of Equally Spaced, Triangularly Shaped Absorption Lines

Shown in Figure 4-3 are two interferograms. The top one (A) is the interferogram of the spectrum with no absorption lines, and the lower one (B) is the interferogram with the absorption lines. The features, called signatures, located at $(\Delta\sigma)^{-1}$ intervals in the B

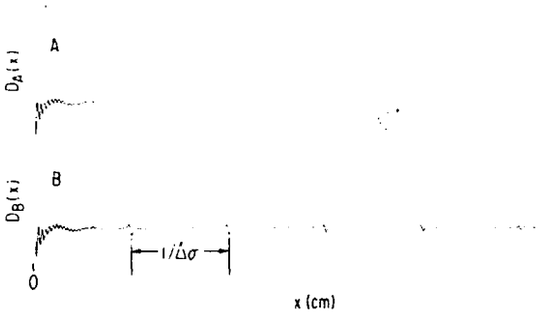


Figure 4-3. Interferograms of Spectrum Shown in Figure 4-2, Without (A) and With (B) the Absorption Lines Present

interferogram are due to the absorption lines. The first signature is somewhat less than 1/25 (that is, ~4 percent) of the central maximum at $x=0$, and although it is not obvious from the figure, the central maximum of the B interferogram is about 4 percent smaller than the central maximum in the A interferogram. Comparison of these two interferograms illustrates the dynamic range problem. Most of the information about the absorption lines is contained in the regions around the signatures. The features are quite small in comparison with the central maximum and become smaller as a function of optical path difference. Later it will be shown that the decreasing amplitude of the signatures is due to the finite width of the lines. Now suppose the detection and amplification system of the interferometer has a dynamic range of 1000, that is, it can discriminate between 1000 gray levels. If the central maximum is made equal to 1000 units, then the first signature would be less than 40 units variation about a 500 unit constant level. This is an undesirable situation, since at best only 4 percent of the dynamic range capability is being used to measure the important information about the absorbing sample.

The example considered above represents not a worst case, but rather a best case as far as the dynamic range problem is concerned. The absorption line positions in the example spectrum (Figure 4-1) are related harmonically. That is, the Mth line is at $M\sigma_0$, then the $(M+1)$ line is at $(M+1)\sigma_0$. Each line will subtract (absorption process) a cosine varying increment from the interferogram and the arguments of the cosine terms are related harmonically. In the neighborhood of $x=M(\Delta\sigma)^{-1}$ (M an integer), the cosine arguments will be integer multiples of 2π and a large variation in the interferogram will be seen. Now if the line positions are not harmonically or near harmonically related there will be, in general, no signatures and the effect of the cosine terms (one for each line) will be spread throughout the interferogram and not localized. In general then, the variation of the interferogram about its average value will be much reduced (but spread out through more of the interferogram) and the dynamic range problem will be much more severe.

The problem may be examined from the fundamental formula of Fourier spectroscopy:

$$D(x)=\int_{\sigma_1}^{\sigma_2} B(\sigma)(1+\cos 2\pi\sigma x)\, d\sigma. \qquad (4\text{-}2)$$

Now for absorption spectra $B(\sigma)$ may be expressed,

$$B(\sigma)=B_0(\sigma)\exp[-k(\sigma)l], \qquad (4\text{-}3)$$

where $B_0(\sigma)$ is the spectral distribution of polychromatic radiation with no absorbing sample present, $k(\sigma)$ is the absorption coefficient and l is the absorbing path. Expanding the exponential in a power series

$$B(\sigma)=B_0(\sigma)\left[1+\sum_{m=1}^{\infty}\frac{[-k(\sigma)l]^m}{m!}\right]. \qquad (4\text{-}4)$$

So,

$$D(x)=\int B_0(\sigma)(1+\cos 2\pi\sigma x)\, d\sigma + \int B_0(\sigma)\left[\sum_{m=1}^{\infty}\frac{[-k(\sigma)l]^m}{m!}\right](1+\cos 2\pi\sigma x)\, d\sigma. \qquad (4\text{-}5)$$

The first integral is the interferogram of the source spectrum with no absorption. The second integral must therefore be the modification of the source interferogram due to absorption. Evaluation of this equation at $x=0$ will give an estimate of the magnitudes of these two terms;

$$D(0)=2\int_{\sigma_1}^{\sigma_2} B_0(\sigma)\, d\sigma+2\int_{\sigma_1}^{\sigma_2} B_0(\sigma)\left[\sum_{m=1}^{\infty}\frac{[-k(\sigma)l]^m}{m!}\right]d\sigma. \qquad (4\text{-}6)$$

The first term is the height of the central maximum with no absorption, and therefore the second term is the reduction of the central maximum when absorption is present. If all the absorption lines absorb r percent of the total $B(\sigma)$, then the central maximum will be reduced by r percent. Since this is the largest value each of these terms may take for any x, it follows that the variation of the interferogram due to the absorption lines cannot be larger, and in general will be much smaller, than r percent.

When wave number measurements are to be made in absorption spectra, narrow or sharp lines are desirable so that their centers may be located accurately. Sharp lines mean small absorption and thus small variation of the interferogram about its average value away from $x=0$, and the detection of these variations will then require large dynamic range.

Equation (4-5) suggests a partial solution to the dynamic range problem. The first integral is the source interferogram with no absorption, and is very large, in general, in comparison to the second term. Therefore, if we can subtract off the source interferogram at the detection stage and amplify the remaining term only, much of the dynamic range problem will be alleviated.

4–3 INSTRUMENTATION FOR A DOUBLE-BEAM DIFFERENCING INTERFEROMETER

An optical solution to this problem is a double-beam differencing interferometer.[2] Essentially it operates on the principle of looking alternately at a reference and a sample beam. If the beams are well matched and there is no absorbing sample present in the sample beam, then the detector will see essentially a DC signal which will not be amplified by an AC amplifier. Absorption in the sample beam will lead to an imbalance and thus an AC signal which will be amplified by the AC amplifier.

The optical diagram of an interferometer of the lamellar grating type, which can be operated in a double-beam differencing mode, is shown in Figure 4-4.

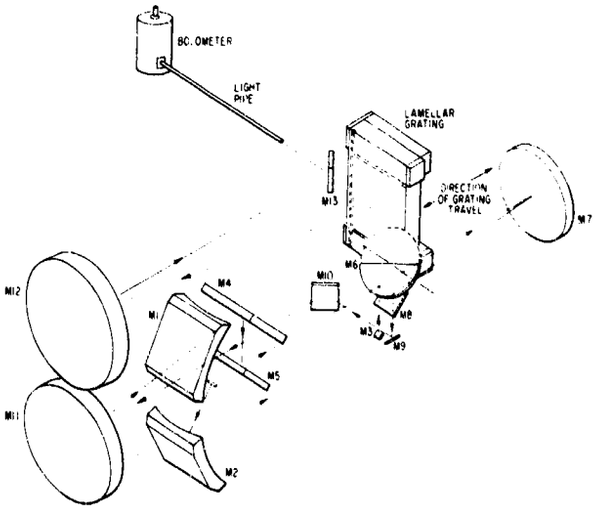


Figure 4-4. Optical Diagram of a Double-beam Differencing, Lamellar Grating Interferometer

Two beams are caught and focused by ellipsoidal mirrors M_1 and M_2. The top or sample beam is folded by flats M_4 and M_5. The focus for this beam is between M_5 and M_6, the chopper. The lower beam, the reference beam, is focused on flat M_3 which folds the beam up to the chopper, M_6. When the chopper is rotated, the sample and reference beams are alternately passed (sample beam) or reflected (reference beam) to the spherical mirror M_7. The rest of the optics are fairly straightforward and will not be discussed further here.[2] Shown in Figure 4-5

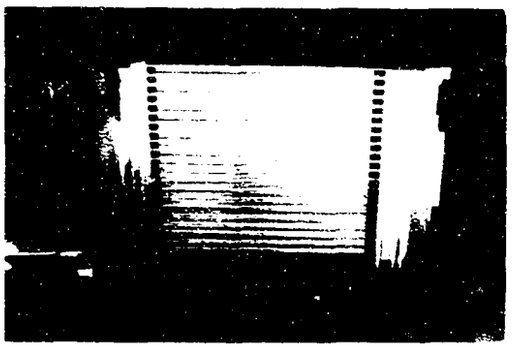


Figure 4-5. A Lamellar Grating of Aluminum Construction. The useful area of the grating is ~30.5 cm×32.8 cm

is a photograph of the lamellar grating now used in this instrument. It is made of aluminum, while the previous lamellar grating was made of pyrex glass.

As an example, the interferogram of the pure rotational spectrum of CO was examined in the double and single beam mode of operation. Carbon monoxide was selected since it is a diatomic molecule and therefore its rotational spectrum will consist of very nearly equally spaced lines. Thus CO will exhibit signatures in its pure rotational interferogram.

Shown in Figure 4-6 are the first, fifth and tenth signatures of CO, single and double beam, for an absorbing path of 15 cm at a pressure of 780 torr. The top curves were run at an amplification of ~20 times that of the lower curves. The information in the upper curves is of course present in the lower curves. To extract the information from the single beam run would however require a vastly larger dynamic range than is necessary for the double-beam run.

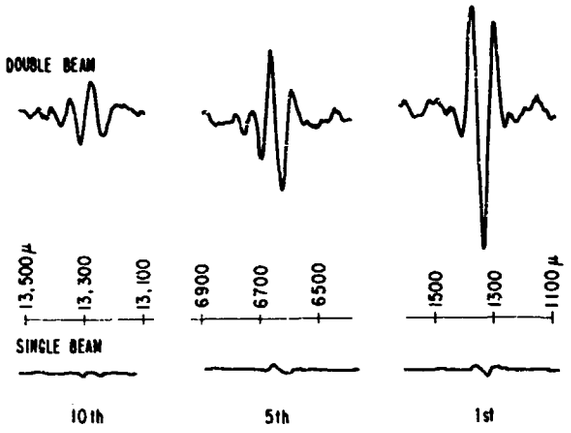


Figure 4-6. Single and Double-beam Signatures (the 1st, 5th, and 10th) from the Pure Rotational Interferogram of CO. $L = 15$ cm, $P = 780$ cm

4–4 PROBLEMS IN DOUBLE-BEAM DIFFERENCING

Double-beam differencing presents some problems. Shown in Figure 4-7 is a double-beam interferogram,

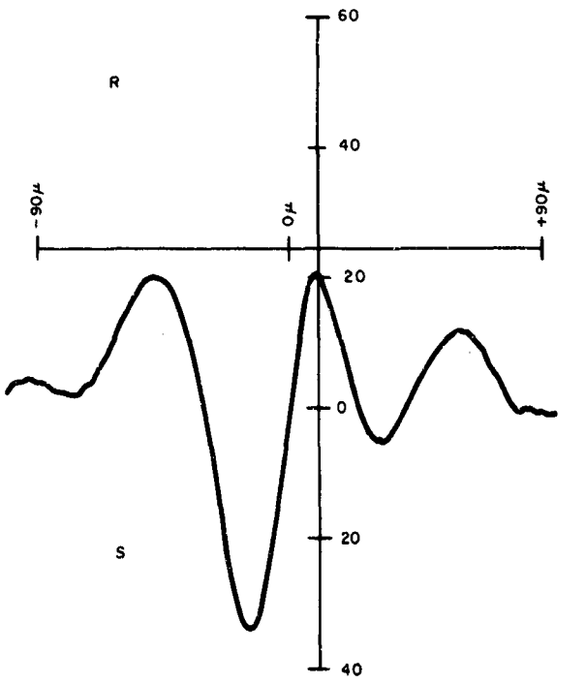


Figure 4-7. The Region About the Origin ($x = 0$), Double Beam. The instrument is misaligned and the spectral distribution of energy in the reference and sample beams are not equal

about $x=0$ when there is misalignment of the instrument and unequal filtering. If the beams had been matched perfectly, the curve should be a straight line having the value zero for all x. In Figure 4-7, the region R (above the zero) is where the reference beam is stronger than the sample beam, the region S (below the zero) is where the sample beam is stronger than the reference beam. The interferogram is not symmetrical about zero, as it should be, and thus misalignment of the instrument is indicated. Shown in Figure 4-8 is a double-beam interferogram about the origin with the misalignment largely corrected, but there is still some imbalance between the beams. The curve is now very symmetrical about the origin, indicating the alignment is now very good. There is residual imbalance, due mainly to a slightly different spectral distribution of radiation in the beams. The matching is really quite good, in that if a single-beam interferogram were run at the same amplification, its central maximum (peak to minimum) would be ~100 times larger than the maximum to minimum of the curve shown in Figure 4-8. Unbalanced beams and misalignment are most apparent in the neighborhood of the origin ($x=0$). When x is greater than a few hundred microns, the imbalance is not detectable even at the highest usable amplification (that is, to within the overall noise level of the system).

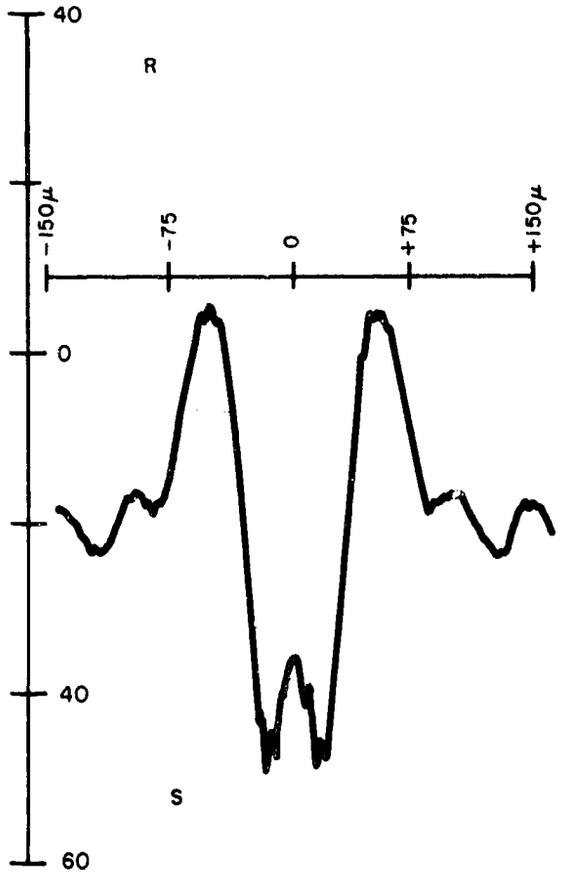


Figure 4-8. The Region About the Origin ($x=0$), Double Beam. The misalignment is now largely corrected, but the spectral distribution of energy in the reference and sample beams are not equal

One of the reasons for the spectral imbalance between the two beams is the envelope of the high pressure mercury arc lamp. The quartz envelopes of these lamps are uniform enough to produce channel spectra in the far infrared, but not uniform enough to produce equal channel spectra from all parts of the arc envelope. The two beams of the instrumentation described here pass through different parts of the quartz envelope. Exact cancellation in double beaming is not achieved and is in part responsible for the spectral imbalance.

The mercury arcs used are also unstable. Fluctuations in arc intensity are small, but not negligible, and spatial movements of the arc on the electrodes occur. In single-beam operation, they appear quite small ($\lesssim\frac{1}{4}$ percent of the total signal) but their behavior is quite noticeable in double-beam operation, and we believe that these fluctuations are responsible for more noise in the derived spectrum than the detector noise. This problem was partially solved by using a bare tungsten filament (geometry is a coiled coil) as a source. The stability (long and short term) can be made very high, and since the source is operated in vacuum, the envelope can be removed eliminating this source of channel spectra. When this lamp is operated at a color temperature of ~2700° K, it is superior to the mercury arc for $\sigma \geq 90\ \mathrm{cm}^{-1}$, about equal in the region $60 \leq \sigma \leq 90\ \mathrm{cm}^{-1}$ and inferior below $\sigma \sim 50$ to $60\ \mathrm{cm}^{-1}$.

Finally, note that a price is paid in double-beam differencing in the form of information lost. The difference between a reference interferogram and a sample interferogram is measured interferometrically. The difference between the reference spectrum and the sample spectrum (which looks like an emission spectrum) is obtained spectrally. The information lost is the 100 percent absorbing curve, that is, where a 100 percent absorbing line would go as a function of σ. Ideally, subsequent and/or previous single-beam source runs (which in general would be quite short since the interferogram is essentially flat after a few millimeters of optical path difference has been attained) would provide such a curve if the change in amplification is accurately known. Although this possibility has not been discussed to any great extent, it appears that in practice, the double-beam spectra cannot be matched very precisely with the single-beam spectrum, and therefore, accurate intensity measurements cannot be made.

4-5 GENERAL PROBLEMS

Now consider some problems encountered with the instrumentation discussed that are not peculiar to double-beam differencing. The problem of phase errors is considered first. An ideal interferometer

gives $D(x)$ as,

$$D(x)=\int_{\sigma_1}^{\sigma_2} B(\sigma)(1+\cos 2\pi\sigma x)\, d\sigma. \qquad (4\text{-}7)$$

In practice, interferometers are never ideal and what is obtained may be written as,

$$D^1(x)=\int_{\sigma_1}^{\sigma_2} B(\sigma)[1+\cos\{2\pi x\sigma+\varphi(\sigma)\}]\, d\sigma, \qquad (4\text{-}8)$$

where $\varphi(\sigma)$ is the phase error. In practice, $\varphi(\sigma)$ may be determined experimentally from a short two-sided interferogram. The prescription is given by Forman *et al.*[3] and can be determined quite easily. A somewhat poor result is shown in Figure 4-9. The open

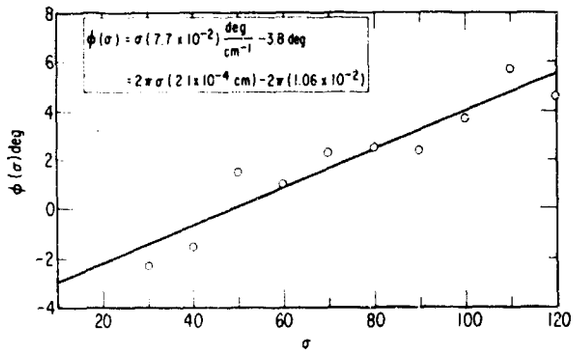


Figure 4-9. The Phase Error of an Interferogram

circles are the experimental points. From the scatter of the points it is apparent only a straight line fit is required. The equation of the least-squares fitted straight line is given at the top of the figure. As written on the second line, the constant multiplying $2\pi\sigma$ is $2.1\mu m$ and is the error made in locating the point $x=0$ in the interferogram. Typically this point can be located somewhat more accurately. The constant term represents a constant phase change and its cause is not well understood. It is fairly small and can be corrected quite easily.

Another problem is that the average value of the interferogram away from $x=0$ does not have a constant value as a function of x. This is illustrated in Figure 4-10, where $\overline{D}(x)$ is plotted as a function of x. This is caused by the shadowing that takes place in the lamellar grating. Such a long term drift will introduce spurious fluctuations of intensity in the very low wavenumber region, since a long term drift of this type corresponds largely to low frequency Fourier components. This type of drift will also

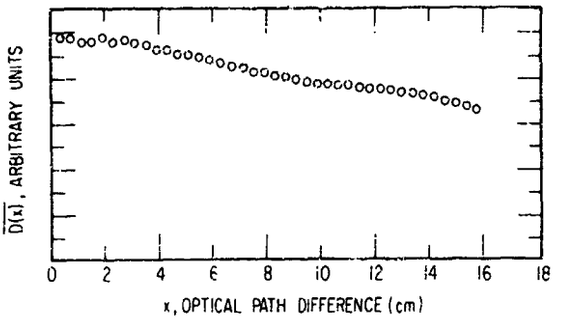


Figure 4-10. Typical Drift of the Average Value of an Interferogram as a Function of Optical Path Difference

broaden spectral lines in any region of the spectra and correction should be applied when line shape measurements are made.

In molecular spectroscopy, the wavenumber position of the absorption lines yields information concerning the geometric and dynamic structural parameters of the molecule being studied. Thus, one of the more important measurements is the absolute position of absorption lines. One of the pure rotational lines in the spectrum of CO as determined from a double-beam interferogram is shown in Figure 4-11. The

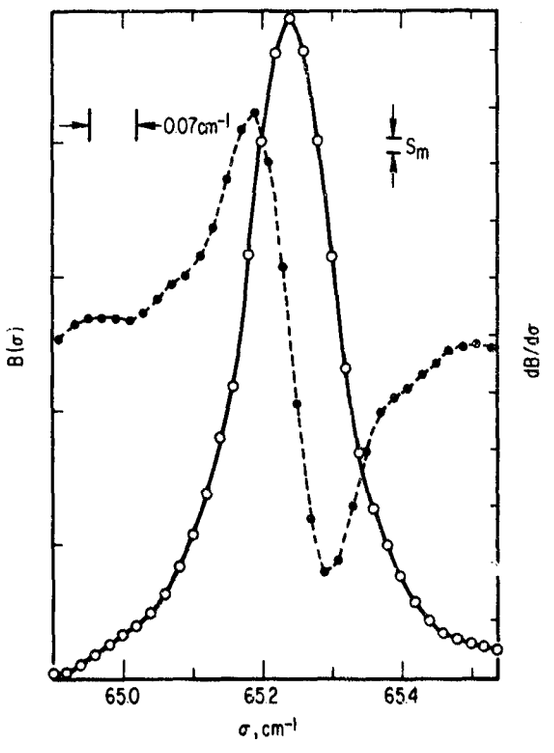


Figure 4-11. The $J=17\leftarrow16$ Pure Rotational Line of CO Derived From an Interferogram Obtained with a Pressure of 608 torr, an Absorbing Path of 15.2 cm, at Room Temperature. The maximum optical path difference attained was 8.7 cm

absorbing path of the gas sample was 15.2 cm and the pressure was 608 torr. The open circles are the output intensity points (absorption goes up) from the computer printout. The line is very symmetric. In order to examine the symmetry more closely, the first derivative of this curve is also shown (solid circles, scale to the right). If the line were perfectly symmetric, the maximum of the first derivative should be equal numerically to the minimum of the first derivative, but of opposite sign. In Figure 4-11, twice the standard derivation of the noise S_m (right) and the full width at half maximum intensity of the scanning function of an ideal two-beam interferometer are shown.

Typical results from good spectra from the instrument (that is, when the phase errors are small) have the maximum and minimum of the first derivative curve of the line shape differing by only a few percent. Also, the maximum and the minimum of the first derivative curve should be equally spaced about the center of the line. A typical set of data is shown in Table 4-1.

Table 4-1. Location of the First Derivative Extrema Relative to the Line Centers for the Pure Rotational Lines of CO*

J_L†	$(\Delta\sigma_D)_-$	$(\Delta\sigma_D)_+$	$[(\Delta\sigma_D)_- - (\Delta\sigma_D)_+] = \delta$
6	0.0477	0.0441	+0.0036
7	0.0498	0.0562	−0.0064
8	0.0570	0.0539	+0.0031
9	0.0649	0.0618	+0.0031
10	0.0590	0.0593	−0.0003
11	0.0649	0.0669	−0.0020
12	0.0593	0.0567	+0.0026
13	0.0611	0.0625	−0.0015
14	0.0615	0.0589	+0.0026
15	0.0599	0.0603	−0.0004
16	0.0531	0.0550	−0.0019
17	0.0502	0.0526	−0.0024
18	0.0539	0.0525	+0.0014
19	0.0496	0.0471	+0.0025
20	0.0516	0.0527	−0.0011

* The absorbing path length was 15 cm and the pressure was 608 torr.
† J_L is the total angular momentum quantum number of the lower state of the transition. The quantity $(\Delta\sigma_D)_-$ is the wavenumber distance from the first derivative maximum to the line center and $(\Delta\sigma_D)_+$ is the wavenumber distance from the line center to the first derivative minimum.

The root-mean-square value of the δs is 2.8×10^{-3} cm^{-1}. This number is considered to be typical of the accuracy to which line centers of well defined isolated lines can be located. Absolute accuracy is discussed next.

The rotational constants of CO have been very accurately determined from careful microwave[4] and near infrared[5] studies. It is possible to predict the pure rotational spectrum lines of CO to better than 3×10^{-4} cm^{-1} or about ten times better than expected. The pure rotational spectrum of CO was examined under conditions of high resolution over twenty times. Figure 4-12 shows a comparison

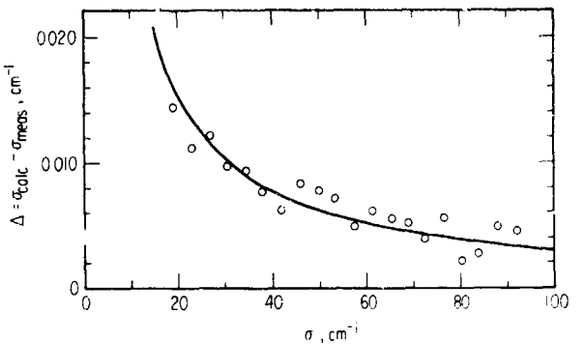


Figure 4-12. The difference Δ, Between the Calculated and Measured Pure Rotational Line Positions for CO. The open circles are the differences for the $J=5\leftarrow4$ through the $J=24\leftarrow23$ transitions. The curve is $\Delta=(0.31\,\sigma^{-1})$

between the far infrared observed, and the calculated positions (from microwave and near infrared data) of the pure rotational lines of CO. Plotted as a function of σ is the difference σ_{calc} (from near infrared and microwave data) $-\sigma_m$ (measured, far infrared) for twenty pure rotational lines of CO observed in the region 20 to 90 cm^{-1}. The curve, which fits the points to better than $\pm3\times10^{-3}$ cm^{-1}, is shown as the solid curve and is of the form $f(\sigma)\propto\sigma^{-1}$. A correction curve of this form is peculiar for spectra measured interferometrically. The types of errors usually made in interferometric measurements are reflected in errors and distortions in the spectra which become larger as σ becomes larger. This wavenumber correction curve has exactly the opposite behavior. The correction is largest for small wavenumbers, becoming smaller for larger wavenumbers.

Such a correction curve can be understood if it is assumed that the lamellar grating acts as a waveguide. Even though the waveguide walls are pyrex glass, the reflectivity of this material (at the angles of incidence the far infrared radiation would encounter in the lamellar grating) is probably high enough so that the cavities formed by the two sets of facets act like waveguides. The phase velocity of the electromagnetic wave inside the lamellar grating will depend on the spacing of the facets, the polarization of the electric field vector and the wavelength of the radiation. When this is taken into account, there will be an apparent wavenumber shift of measured lines and the shift will be proportional to σ^{-1}. There will also be an apparent broadening of the spectral lines, and if line shape measurements are to be made, a correction should be applied.

4-6 TYPICAL RESULTS

The raw data of a two-beam interferometer ar. called an interferogram. As mentioned previously, the interferogram is related to the spectrum by a Fourier transformation. A few interferograms (simple and complicated) will be considered in this section. Real time information on the progress of the interferometric measurements, and therefore also on the spectroscopic measurements can be obtained.

Portions of a double-beam interferogram of CO obtained in a ~15 cm cell at a pressure of 202 torr are shown in Figure 4-13. Qualitative interpretation of

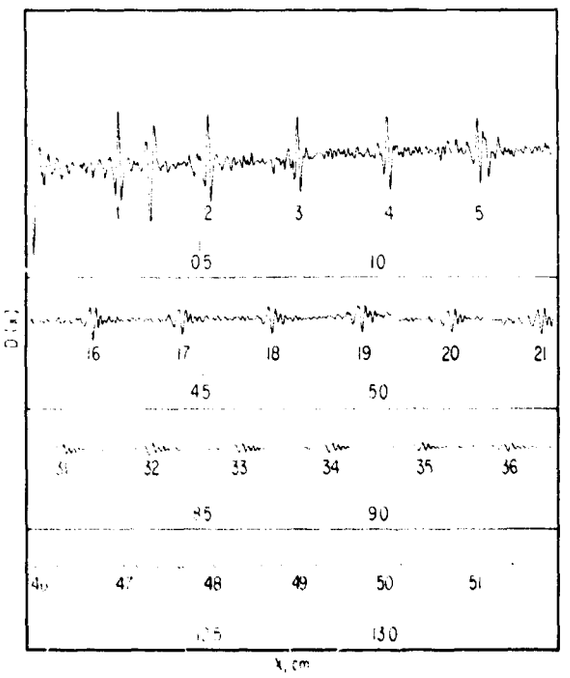


Figure 4-13. Portions of the Pure Rotational Interferogram of CO

the interferogram is straightforward and a number of observations can be made directly concerning the spectrum from the interferogram. Since CO is a diatomic molecule, its spectrum will consist of nearly equally spaced lines. Equally spaced lines in the spectrum lead to equally spaced features in the interferogram, as seen previously. The signatures are easily seen in the interferogram. Signatures are not uniform in shape, but rather are distorting as one goes to larger optical path differences. This happens because the lines in the CO spectrum are not exactly equally spaced, rather converge slightly as a function of rotational quantum number. (This in turn is due to the nonrigidity of the molecule.) The signatures are decreasing in amplitude as a function of optical path difference, over and above the apparent loss in amplitude due to the distortion of the signatures. This is due to the finite width of the lines and will be discussed in more detail later. There is an extra signature at $x \sim 4$ mm due to the channel spectrum of the lamp envelope of the source, previously mentioned.

Figure 4-14 shows the spectrum derived from this interferogram in the region $\sigma = 20$ to 86 cm^{-1}. The CO lines are the strong lines located ~3.8 cm^{-1} apart.

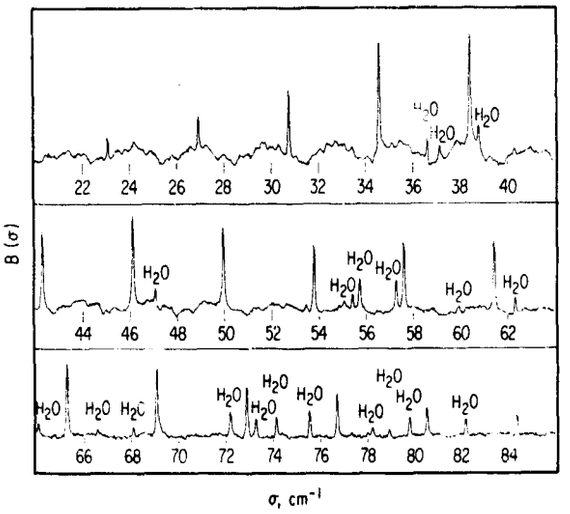


Figure 4-14. The Pure Rotational Spectrum of CO Between 20 and 86 cm^{-1}

The weaker lines are water vapor lines. The H_2O molecule is a very strong absorber in the far infrared and its presence in even very small quantities will be reflected in the spectrum by the presence of its strongest lines. The cosinusoidal variation of the base line, especially prominent at low wavenumbers, is the channel spectrum of the quartz envelope of the high pressure mercury source lamp. Absorption in the figure goes up, but since the double-beam differencing mode of operation gives no information about the 100 percent absorbing curve, comparison of the intensity of lines in widely separated wavenumber regions could be erroneous.

Figure 4-15 shows a portion of the far infrared double-beam interferogram of nitric oxide.[6] This interferogram is somewhat more complicated than the CO interferogram. First, there are actually two sets of signatures (denoted by the a and b sequence) that are well resolved interferometrically beyond about $x = 30$ mm. It follows then that the spectrum will consist of two sets of almost uniformly spaced lines, the spacing for the a sequence being larger than that for the b sequence. Secondly, the a sequence is distorting (smearing out) more rapidly than the b sequence. Therefore, the lines corresponding to the

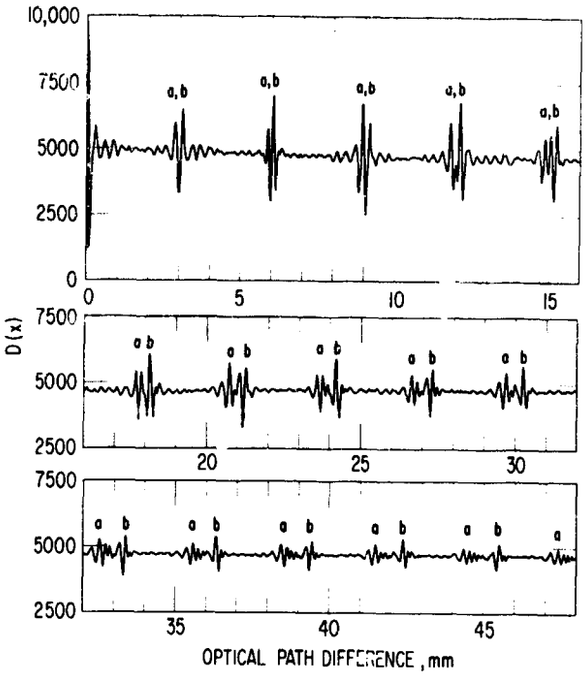


Figure 4-15. A Portion of the Pure Rotational Interferogram of NO

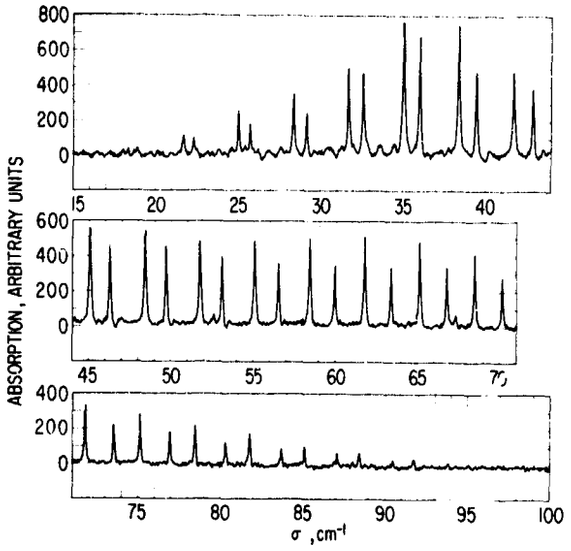


Figure 4-16. The Pure Rotational Spectrum of NO in the Region 15 to 100 cm^{-1}

a sequence of signatures will show a greater covergence (because of centrifugal distortion and other effects) than the lines corresponding to a b sequence. Finally, adjacent signatures in each set are inverted copies of each other (note, for instance, the b signature at $x \sim 39$ mm and at $x \sim 42$ mm). This implies that the quantum number in the energy level expression is nonintegral. In fact, since the adjacent signatures are completely inverted with respect to each other, it follows that the quantum number J must be half integer.

It is well known that the ground state of the NO molecule is a $^2\pi$ state which is split into $^2\pi_{1/2}$ and $^2\pi_{3/2}$ states by spin orbit interaction. The two states are separated by $\sim 120\ cm^{-1}$ and thus two sets of rotational lines are expected, one for each state. Furthermore, the overall rotational quantum number is known to be odd multiples of one half, which is directly attributable to the odd number of electrons in the NO molecule.

Figure 4-16 shows the spectrum derived from this interferogram. That there are two sets of lines is obvious and the analysis shows that one sequence of lines converges faster than the other sequence, and the rotational quantum number J must be half integral.

Figure 4-17 shows a very complicated interferogram of water vapor in a 15 cm cell at a pressure of 10 torr. The interpretation of individual features seems hopeless, however, the overall behavior of the interferogram does provide valuable information for interpreting the derived spectrum.

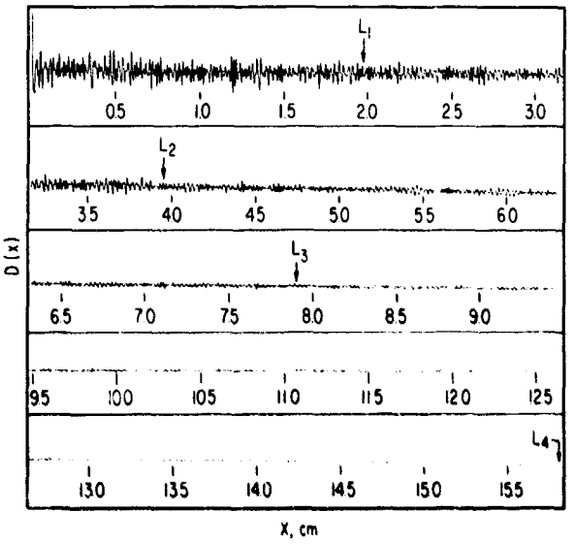


Figure 4-17. The Pure Rotational Interferogram of Water Vapor

Figure 4-18 shows the spectrum derived from this interferogram. It is the pure rotational spectrum of the H_2O molecule between 15 and 125 cm^{-1}. The water vapor molecule is a very asymmetric top, and thus is expected to have a very complicated pure rotational spectrum. Each line in this spectrum has been identified with a pure rotational transition and the analysis of this spectrum has led to very accurate rotational constants of the H_2O molecule.[7,8]

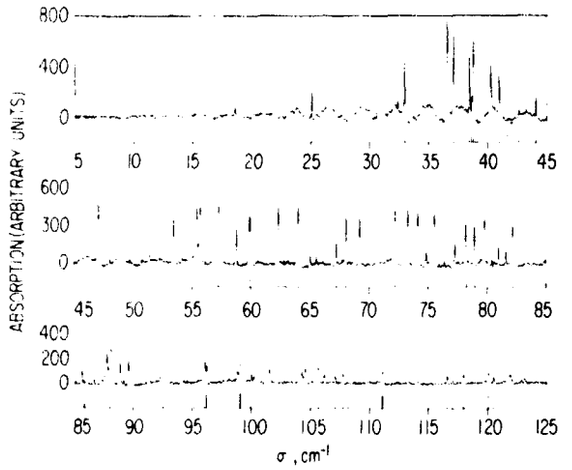


Figure 4-18. The Pure Rotational Spectrum of Water Vapor Between 5 and 125 cm^{-1}

4-7 FINITE WIDTH OF LINES

The scanning function of a spectroscopic instrument may be thought of as the output spectrum when an infinitely narrow line (δ function) is the input spectrum. For an ideal two-beam interferometer, the interferogram of an infinitely narrow line will be a cosine curve of constant amplitude as a function of optical path difference. The output spectrum from an interferogram which is a cosine curve, truncated at $x = L$, is shown in Figure 4-19 and is the scanning function of an ideal two-beam interferometer. The curve is the familiar sinc function centered about σ_0, the position of the δ function line input spectrum. The wavenumber scale is in units of $(L)^{-1}$, where L is the maximum optical path difference attained in the interferogram. If all lines looked like this, Fourier spectroscopy would lose much of its attractiveness. In particular, increasing L will bring the extraneous sidelobes in closer to the line at σ_0, but the intensity ratio of the sidelobe intensity to the main line intensity would remain constant. The confusion that could result from two or more closely spaced lines is not difficult to imagine.

The more physical situation of lines with finite widths is not as serious, providing the interferogram can be measured to where the width of the line is comparable to the resolution attained. Figure 4-20 shows the familiar Lorentz line shape. The curve has its maximum value at $\sigma = \sigma_0$ (arbitrarily normalized to unity at σ_0) and ϵ is the half width at half-maximum intensity.

The interferogram such a line would present is shown in Figure 4-21. The interferogram is a damped cosine curve, and in particular for a Lorentz line, the argument of the damping exponential function is proportional to the product of the line width parameter ϵ and the optical path difference x. As mentioned

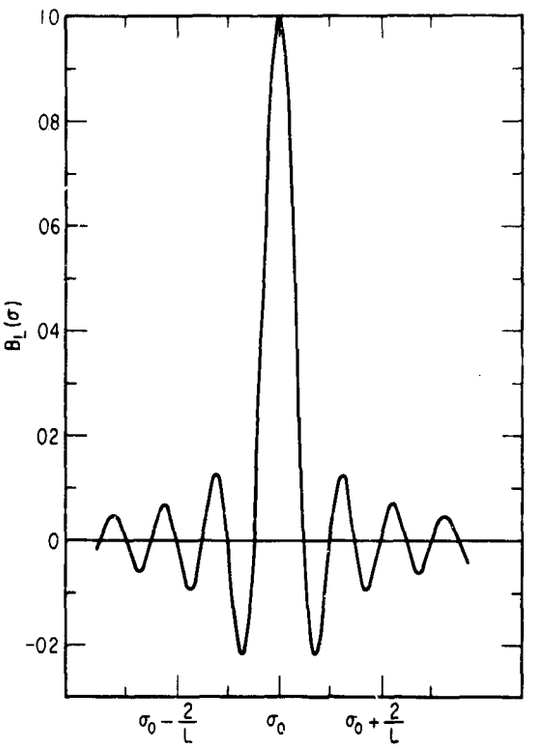


Figure 4-19. The Scanning Function of an Ideal Two-beam Interferometer

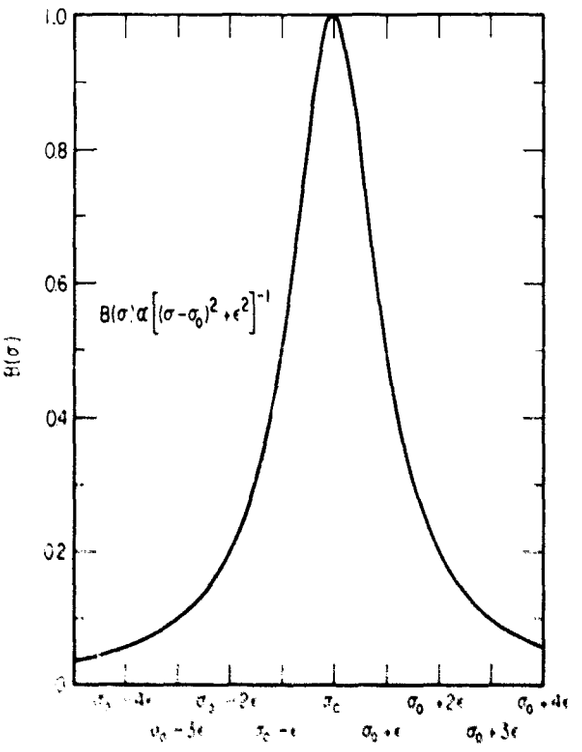


Figure 4-20 The Lorentz Line Shape. σ_0 is the resonant wavenumber and ϵ is the halfwidth at half-maximum intensity of the line

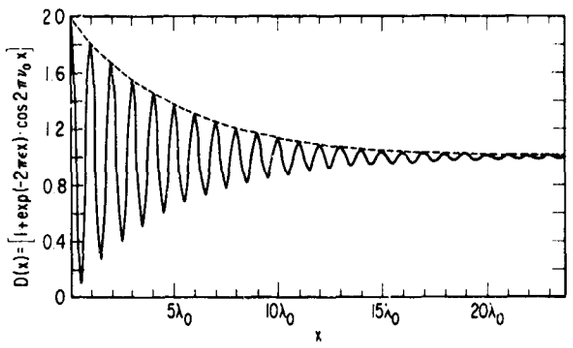


Figure 4-21. The Interferogram of a Single Lorentz Line

earlier, the finite width of the lines in the CO spectrum was responsible for the decay in amplitude of the signatures in the interferogram.

Obviously, if $D(x)$ shown in Figure 4-21 could be measured out to infinity and could be transformed properly, the original Lorentz line would be back. Realistically, of course, the interferogram will be truncated at some finite value of x, and therefore the loss of the information beyond $x=L$ will introduce some distortion into the line shape.[9] Figure 4-22 shows two lines; the solid curve is the original Lorentz line, the line that would be obtained from an infinitely long, noiseless interferogram as shown in Figure 4-21. The dashed curve is the line shape that would be obtained if the interferogram were terminated when the product $L\epsilon=0.200$. The sidelobes are still quite apparent, but are smaller than for the δ function line. In fact, the intensity of the sidelobes is now a function of L, the maximum optical path difference attained in the interferogram. The inset in Figure 4-22 shows the difference between the two line shapes as a function of σ. The sidelobe intensity is defined as the difference between the second maximum (at $\sigma\sim\sigma_0+4\epsilon$ in Figure 4-22) and the first minimum (at $\sigma\sim\sigma_0+1.5\epsilon$ in Figure 4-22). This is compared to the intensity of the line at $\sigma=\sigma_0$. The ratio of the line height to the sidelobe height, as a function of $L\epsilon$ is shown in Figure 4-23 for $0\leq L\epsilon\leq 0.9$.

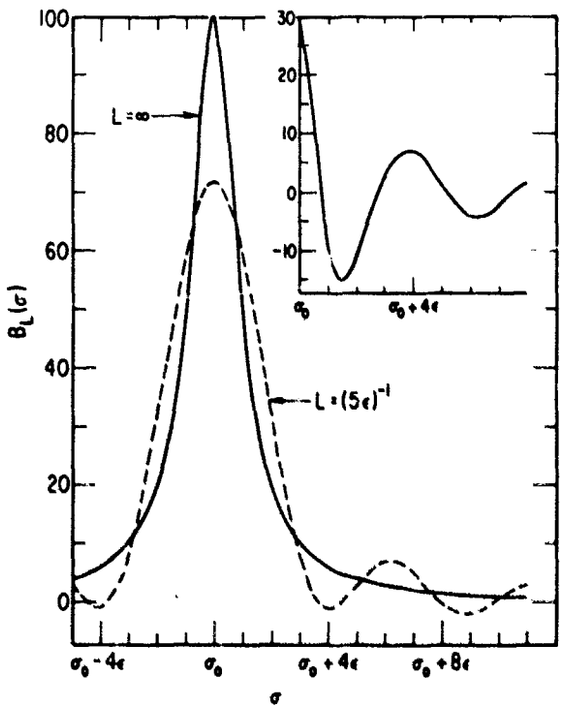


Figure 4-22. Lorentz Line Shape as Seen by an Ideal Two-beam Interferometer. The solid curve is the line that would be obtained if $L\rightarrow\infty$. The dashed curve is the line obtained if $L=(5\epsilon)^{-1}$. The inset shows the difference between the two line shapes

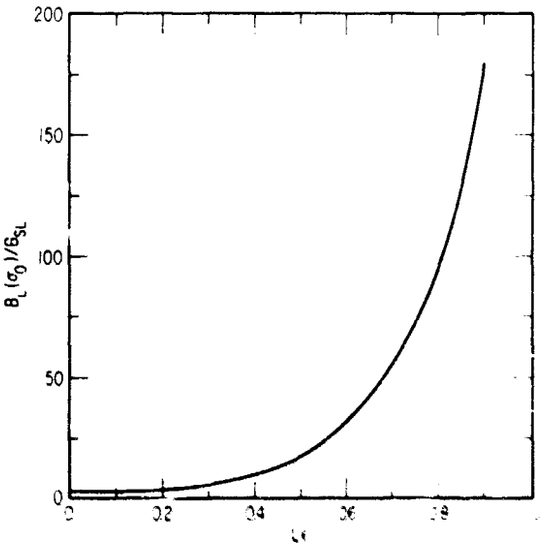


Figure 4-23. Plot of the Ratio of Line Intensity $B_L(\sigma_0)$ to Sidelobe Intensity, B_{SL}, as a Function of $L\epsilon$. L is the maximum optical path difference attained in the interferogram and ϵ is the line width parameter of the Lorentz line

For $L\epsilon$ small, that is, the resolution of the interferogram being grossly inferior to the line width parameter of the line, the sidelobe intensity will be quite prominent. But as the product $L\epsilon$ is made larger, the ratio of the line intensity to the sidelobe intensity increases dramatically. For instance, when $L\epsilon\sim 0.2$ the ratio of line intensity to sidelobe intensity is ~3.4, whereas when $L\epsilon\sim 0.8$, the ratio of intensities is ~100. By quadrupling the resolution (that is, L for constant ϵ), the presence of the sidelobes has been reduced by a factor ~30.

When the interferogram has to be terminated before the sidelobe intensity becomes negligible it will be

advisable to artificially broaden the lines by a mathematical operation on the interferogram called apodization. Essentially, apodization is multiplying the interferogram (less the constant term) by a monotonically decreasing function of x. Apodization is easily done by computers and rather extensive treatments of this process exist in the literature,[10] so the mechanics of apodization will not be further pursued here.

The above discussion is strictly quantitative only for Lorentz shaped lines. (A similar analysis is, however, possible for Gaussian shaped lines but will not be discussed here.) In absorption spectroscopy the Lorentz line shape is strictly appropriate only to weakly absorbing lines. However, even if the lines are fairly badly saturated (even if the maximum intensity of absorption ~99.5 percent), the line shape is still quantitatively Lorentzian[11] (except that an effective half-width parameter is now defined which is in general larger than the true line width parameter).

To illustrate the difficulties of trying to interpret spectra obtained from interferograms with inadequate resolution, refer to Figure 4-17, the interferogram of the pure rotational spectrum of the water vapor molecule. The interferogram is marked at four points: L_4, L_3, L_2, and L_1 corresponding to maximum resolution, one-half maximum resolution, one-fourth maximum resolution, and one-eighth maximum resolution, respectively. The spectra obtained by terminating the interferogram at each of these points are shown in Figure 4-24. The region between 54 and 84 cm^{-1} is shown. The top spectrum was obtained by terminating the interferogram at L_1 while the bottom spectrum corresponds to terminating the interferogram at L_4. If only the L_1 interferogram spectrum were available, it might be possible to claim that there is a line between the stronger lines at 59 and 60 cm^{-1}, and two lines between the stronger lines at 62 and 64 cm^{-1}. The situation is only slightly better in the L_2 interferogram spectrum. Most of the sidelobes are of smaller intensity, but not completely so. Now the single line between the 68 and 69 cm^{-1} lines has been resolved into three lines, two of which have larger intensities than the original single line. The L_3 interferogram spectrum does show marked improvement, but some sidelobes are still obvious between the 59 and 60 cm^{-1} lines. The L_4 interferogram spectrum appears to be free of sidelobes. There was still some information in the interferogram between L_3 and L_4 as the single line at 60 cm^{-1} is now split into a doublet.

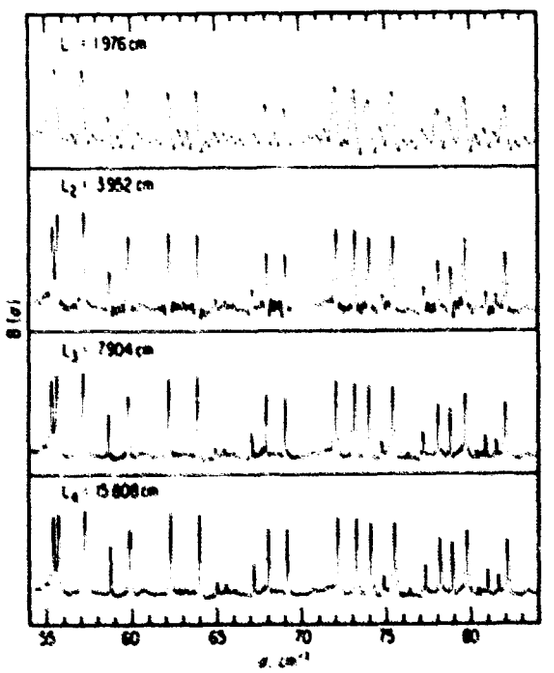


Figure 4-24. The Pure Rotational Spectrum of Water Vapor Between 54 and 84 cm^{-1}. The spectrum is computed from the interferogram shown in Figure 4-17 truncated at x = 1.976 cm (top), 3.952 cm, 7.904 cm, and 15.808 cm (bottom of figure)

Speaking with all the wisdom and authority that hindsight provides, it is quite obvious that if the L_1 or L_2 interferogram spectra were the best one could do, severe apodization should be applied. The L_3 interferogram spectrum is much better, but its interferogram too should have been apodized (perhaps more gently) since the sidelobes are still apparent, and for closely spaced lines the overlapping sidelobes could lead to inaccurate location of the line center. The L_4 interferogram spectrum appears to be free from sidelobes and thus apodization is not required.

As mentioned above, this is hindsight, and of no particular use while obtaining the data, that is, the interferogram. Consider two possible solutions. The first and best is to have a small dedicated computer, or access to a large general purpose computer with time-sharing capabilities so that near real-time transformation of the interferometric data can be accomplished. This solution is bound to be somewhat expensive. Another, and less expensive way, is to try to understand the behavior of the interferogram in a quantitative sense as a function of optical path difference so that an assessment can be made of how much more information can be obtained by going further in optical path difference, or of whether or not sidelobes are going to be a problem.

For simple spectra showing signatures (that is, spectra having periodic or near periodic arrangements of lines), the fall-off of signature amplitude as a function of optical path difference can be watched. In general, the signature amplitude will fall off faster than the effective line width would dictate since in general the lines will not be exactly harmonically spaced.

More complicated interferograms do not appear to be entirely hopeless. In the H_2O interferogram (Figure 4-17), there are no apparent regular structures, such as signatures. The fluctuations about the average value of the interferogram are however seen to decrease as a function of optical path difference x. The average behavior of a water vapor interferogram is shown in Figure 4-25. The natural logarithms of the

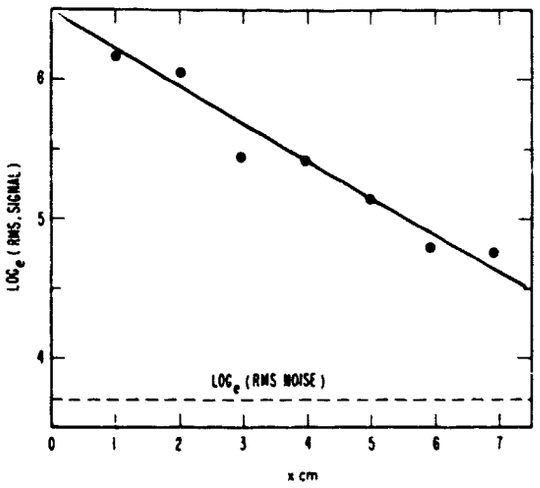


Figure 4-25. Plot of the Natural Logarithm of the Root-mean-square Deviation of the Signal in a Water Vapor Interferogram as a Function of Optical Path Difference

root-mean-square values of the fluctuations of the interferogram, that is, the signal vs optical path difference x are plotted. The points were calculated in the following way. For each point, 40 values of the interferogram centered about the point were averaged. This average value was then subtracted from each of the original 40 values. The root-mean-square deviation was then calculated from these 40 differences and its natural logarithm was plotted as shown in Figure 4-25. A straight line was arbitrarily fitted to these points. A measure of the noise in this interferogram was also determined by measuring 40 values of the interferogram with the interferometer held stationary at a large value of x. The root-mean-square deviation was calculated from the 40 values and the natural logarithm of this value is shown as the straight dashed line in Figure 4-25. Extrapolating the signal line, it intersects the noise line at $x \sim 11$ cm. This is not a prescription to say when the interferogram should be terminated, but only when the signal in the interferogram will become comparable with the noise in the interferogram. How much before or after this point the interferogram should be terminated will depend on many things, including the kind of measurement to be made from the spectrum.

Curves of this type are also useful to determine whether apodization is required, and to what degree. An estimate of the noise in the interferogram leads to an estimate of noise in the spectrum. The slope of the signal curve at larger optical path differences leads to an estimate of the effective line width parameters for the sharpest lines in the spectrum, and since L will be known, Le_m can be evaluated. From the curve given in Figure 4-23, the ratio of line intensity to sidelobe intensity may be obtained. If this number is significantly smaller than the signal to noise in the spectrum for any sharp line, then any line of lesser intensity in the near neighborhood of this line or lines should be viewed with suspicion.

As mentioned previously, mathematical apodization can be thought of as an artificial way of broadening the lines so that the sidelobes do not cause interpretation or measurement difficulties. When possible, it might be better to broaden the lines physically. In absorption spectroscopy of gases, the line width parameter of the absorption lines depends linearly on the pressure except at very low or very high pressures. Quantitatively, increasing the pressure is equivalent to exponentially apodizing the interferogram. Broader lines will also remove more radiation, and thus at least at the beginning of the interferogram, the signal in the interferogram will be larger.

The discussion of simple and complicated interferograms does not encompass the range of interferograms. There are also deceptive interferograms and one is shown in Figure 4-26. This is the inter-

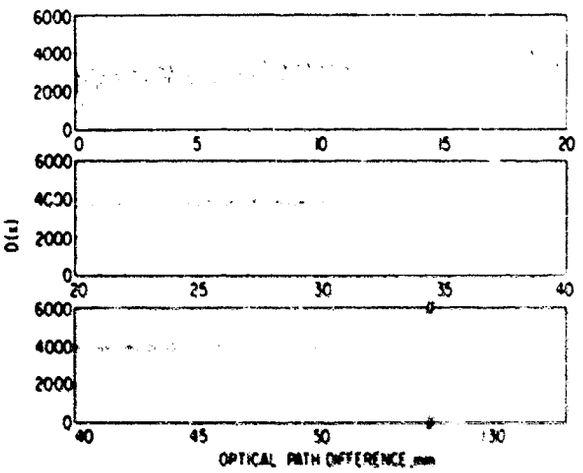


Figure 4-26. The Pure Rotational Interferogram of DCl

ferogram of the rotational spectrum of deuterium chloride (DCl). It is a diatomic molecule and therefore its interferogram shows signature structure, and in this particular interferogram, one can follow the signatures out to $x \sim 25$ mm. In the region around $x \sim 30$ mm, there are fluctuations in the interferogram but no apparent signatures. In the region around $x \sim 35$ mm and larger, the signatures reappear, somewhat more smeared out, but certainly recognizable from ~35 mm and beyond. Spectroscopically, this is due to the fact that there are two isotopic species of DCl, DCl^{35} and DCl^{37}. Each has its own pure rotational spectrum of nearly equally spaced lines, but the spacing is slightly smaller for the DCl^{37} isotope than for the DCl^{35} isotope. Thus there are two sets of signatures which for small x (low resolu-

tion) are in step, but for larger x get out of step and destructively interfere, and for still larger x get back into step, one lobe displaced, and constructively interfere again.

The disappearance and subsequent reappearance of the signatures is indicative of at least some of the isotopic splitting of the pure rotational spectrum of DCl being resolved, and indeed in the transformed interferogram (that is, the spectrum), an extensive series of doublets were resolved.[2,12] Most, if not all of the doublet structure would not have been resolved if the interferogram were terminated when the signature structure disappeared the first time.

4–8 EDITING

In conventional spectroscopy, regions of little or no discrete absorption tell little about the molecule, and these regions are usually not examined extensively. (There are, of course, exceptions.) In interferometry, a narrow class of spectra show the same behavior in the interferogram.[13,14]

A popular and accurate method of determining the real part of the refractive index of reasonably transparent materials is the channel spectrum technique. A plane parallel slab of the material to be investigated is inserted in the sample beam. Part of the incident radiation is transmitted directly through the material but some is internally reflected twice, emerging in the same direction as the single pass radiation. The triply passed radiation suffers a retardation of $2nd$, where n is the real part of the index of refraction and d is the thickness of the sample. Constructive interference takes place if $2nd$ is an integral number of wavelengths, and destructive interference takes place if $2nd$ is an odd number of half integral wavelengths. The spectrum produced is the well known channel spectrum which has the shape of a cosine curve in wavenumber space (neglecting absorption and dispersion). The Fourier cosine transform of the spectrum can be measured interferometrically and the Fourier cosine transform of a cosine function is a δ function. A δ function will be observed only if the cosine curve in wavenumber space is infinitely long. If the spectrum has finite extent, then the δ function will be distorted, as one would expect, into the familiar sinc function. Thus the interferogram of a channel spectrum would be a sinc function in optical path difference space and will be highly localized. The double beam interferogram of a parallel sided piece of fused quartz is shown in the upper left of Figure 4-27. The channel signature, looking much like a sinc function, is quite apparent at $x \sim 0.8$ cm. The spectrum derived from this interferogram is shown at the upper right of the figure. If all the information about the

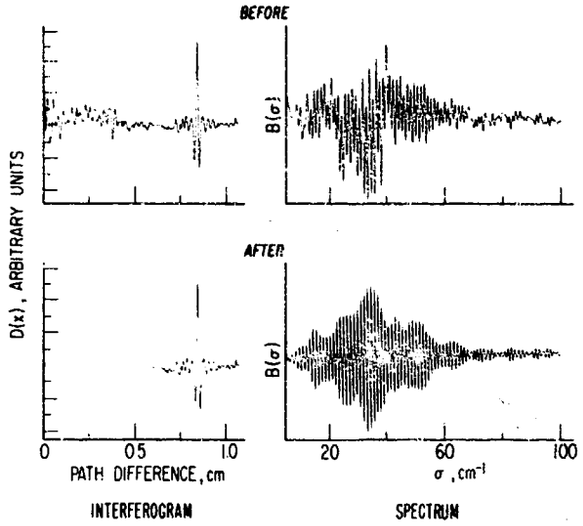


Figure 4-27. The Interferogram and Channel Spectrum of a Plane Parallel Slab of Fused Quartz Before and After Editing

cosine nature of the channel spectrum is concentrated in the signature at $x \sim 0.8$ cm, and the fluctuations in the interferogram outside this region are noise or unimportant, then the interferogram can be edited by replacing the noisy uninformative region with a straight line. The lower left of Figure 4-27 shows the edited interferogram where the region $x \leq 0.6$ cm has been replaced by a straight line. The spectrum derived from this interferogram is shown at the lower right of Figure 4-27. Cosmetically, the channel spectrum looks better, and in fact by editing the interferogram in this manner the low frequency noise (and any real features) have been eliminated. The high frequency noise, of course, remains. In general, editing of this type would be bad interferometric practice even if wide regions of the interferogram were devoid of information. A better procedure would be to skip entirely the regions of no interest and use the available measurement time in the region of interest.

4–9 SUMMARY

Optical double-beam differencing in Fourier spectroscopy has been described and illustrated. The extent to which this technique alleviates the dynamic range problem has been discussed. A number of double-beam interferograms have been shown, and used to illustrate how the behavior of the interferogram is reflected in the derived spectra.

Acknowledgments

Much of the work described in this paper was done in collaboration with Drs. R. T. Hall and C. M. Randall. The technical assistance of Mr. T. E. Mott is also appreciated. This research was conducted under U.S. Air Force Space and Missile Systems Organization (SAMSO) Contract FO 4701–69–C–0066.

References

1. Vanasse, G. A., and Sakai, H. (1967) *Progress in Optics, VI*, E. Wolf, Ed., North-Holland, Amsterdam. p. 263.

2. Hall, R. T., Vrabec, D., and Dowling, J. M. (1966) *Appl. Opt.* **5**: 1147.

3. Forman, M., Steel, W. H., and Vanasse, G. A. (1966) *J. Opt. Soc. Am.* **56**: 59.

4. Jones, G., and Gordy, W. (1964) *Phys. Rev.* **135**: 295.

5. Rank, D., Skorinko, G., Eastman, D., and Wiggins, T. (1960) *J. Mol. Spectry.* **4**: 518.

6. Hall, R. T., and Dowling, J. M. (1966) *J. Chem. Phys.* **45**: 1899.

7. Hall, R. T., and Dowling, J. M. (1969) *J. Chem. Phys.* **47**: 2454.

8. Hall, R. T., and Dowling, J. M. (1970) *J. Chem. Phys.* **52**.

9. Dowling, J. M. (1967) *Appl. Opt.* **6**: 1580.

10. Vanasse, G. A., and Sakai, H. (1967) *Progress in Optics, III*, E. Wolf, Ed., North-Holland, Amsterdam. pp. 284–287.

11. Dowling, J. M. (1969) J.Q.S.R.T. **9**: 1613.

12. Dowling, J. M., and Hall, R. T. (1967) *J. Phys. Radium*, suppl. M° 3–4, Tome 28, C2, 156.

13. Loewenstein, E. V. (1962) *J. Opt. Soc. Am.* **52**: 1023.

14. Randall, C. M., and Rawcliffe, R. D. (1967) *Appl. Opt.* **6**: 1889.

Contents

5. Amplitude Spectroscopy

E. E. Bell
The Laboratory of Molecular Spectroscopy and Infrared Studies, Physics Department, The Ohio State University, Columbus, Ohio 43210

Abstract

Amplitude spectroscopy, which measures the magnitude and phase changes in the electromagnetic wave incident upon a sample, has some characteristics that are quite different from those of power spectroscopy, which measures the power level changes in the incident wave. The relations between these two kinds of spectroscopy are described by the use of the impulse response function of the sample, as measured with asymmetric and symmetric interferograms. Some advantages and disadvantages of amplitude spectroscopy are illustrated by measurements of the transmission and reflection of solids and of the transmission of a gas and a liquid.

5–1 INTRODUCTION

This Conference on Fourier-transform spectroscopy and the literature[1,2] contain a host of discussions of the advantages and remarkable properties of the Fourier transform technique. My purpose is to describe the use of this Fourier transform technique with a Michelson interferometer in the asymmetric mode of operation. This mode of operation has been used with success in two laboratores: The National Physical Laboratory, Teddington, England, by Gebbie and Chamberlain [3–9]; and at The Ohio State University by Sanderson, myself, and our students.[10–16,24] The asymmetric mode of operation provides phase information so that we can achieve an amplitude spectroscopy—the measurement of the amplitude transmittance or reflectance with both the magnitude and the phase determined—in contrast to the normal power spectroscopy—the measurement of the power transmittance or reflectance. For some problems the extra phase information is of no value and amplitude spectroscopy is to be avoided; for other problems the phase information may be the whole purpose, in a measurement of the index of refraction, for example, and then amplitude spectroscopy may be the best measurement choice.*

* A paper by Chamberlain, Gibbs, and Gebbie (1969) *Infrared Physics*, 9: 185, which appeared after the preparation of this paper, describes the use of amplitude spectroscopy in detail.

5–2 RELATIONSHIP BETWEEN A SAMPLE'S RESPONSE FUNCTION AND ITS INTERFEROGRAM

Let us begin our consideration of amplitude spectroscopy by looking at a schematic diagram of a Michelson interferometer that has been used in the asymmetric mode. Figure 5-1 shows the instrument[11]

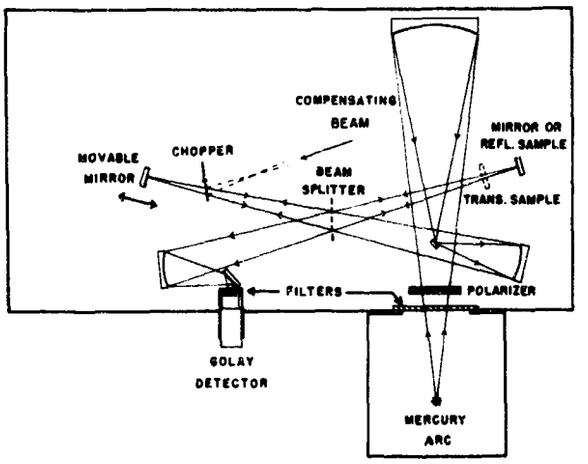


Figure 5-1. The Optical Plan of a Vacuum Far Infrared Michelson Interferometer Used in the Asymmetric Mode for Amplitude Spectroscopy

that was used to obtain the results that illustrate this report. You will note the following features of this particular vacuum far infrared instrument: (1) the sample is placed in one arm of the interferometer; (2) samples can be measured either in transmission or in reflection; (3) the arm without the sample contains the movable mirror for scanning the interferogram; (4) so that small samples can be used, the radiation beam is not collimated in the interferometer arms, on the contrary, the source is imaged on the sample in one arm; (5) the radiation signal is chopped in one arm of the interferometer, and the backs of the chopper blades are used to reflect a compensating signal into the detector path. The compensating signal is adjusted to reduce the chopped detector signal to zero when the instrument is measuring a very large path-difference interferogram signal.

The diagram does not indicate the fact that the transmission measurements are made with the radiation passing only once through the sample. The beam transmitted through the sample does not retrace the incident beam path, but returns to the beam splitter above the plane of the diagram. When making transmission measurements, a corresponding change is made in the other arm of the interferometer, and the detector is moved to receive the signal. The movable mirror is positioned at each measuring point in the interferogram by a simple micrometer screw that is driven by a stepping motor. The interferogram signal is read by an integrating digital voltmeter while the optical path remains fixed at the measuring position. The temperature of the interferometer is kept a few degrees above ambient temperature by a thermostatically controlled electric heating belt wrapped around the vacuum tanks.

A measurement of a sample includes running a background interferogram without the sample, and an interferogram with the sample in place. The effect of introducing a transmission sample into one arm of the interferometer is illustrated in Figure 5-2. The transparent polyethylene sample delayed the passage of the radiation through the sample arm of the interferometer. The optical path of the other arm had to be increased to compensate for this delay in order to locate the strong peak of the "white light position." It is evident that the "white light position" was shifted a distance $D=(n-1)b$, where n is the index of refraction and b is the thickness of the sample. A measurement of the thickness and the shift of the "white light position" is thus a measurement of the index of refraction.

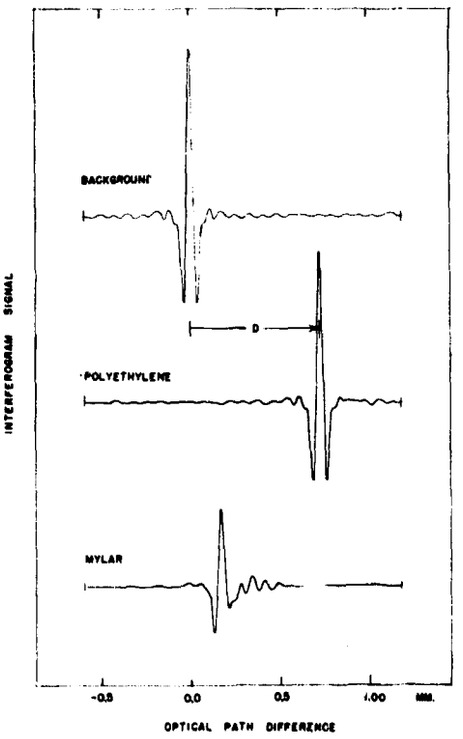


Figure 5-2. The Central Portions of a Background Interferogram, a Polyethylene Sample Interferogram, and a Mylar Sample Interferogram, to Show the Displacement and Distortion of the Central Structure in Amplitude Spectroscopy. The thickness of the polyethylene was 1.510 mm; the thickness of the Mylar was 0.249 mm

The interferogram for the Mylar film shown in Figure 5-2 is distorted into an asymmetric function with no clearly defined "white light position" due to the dispersion of the sample. Because the index of refraction depends upon the wavelength, all wave vector components are not shifted by the same D value. A Fourier analysis of the interferogram gives the phase ϕ for each wave vector σ, so that a value of D, given by $D=\phi/2\pi\sigma$, and an index of refraction can be calculated for each wave vector.

The detector signal in the interferometer is proportional to the time average of the square of the electric field vector resulting from the superposition of the waves from the two arms of the interferometer. This detector signal depends upon the optical path difference x between the arms, and it consists of a constant portion that is independent of x plus an "interference portion" that is dependent upon x. The interferogram is this "interference portion" of the detector signal as a function of the optical path difference and has an x-averaged value of zero. Thus the interferogram is proportional to the time average of a product of the electric field vectors of the wave from the sample arm and the wave from the adjustable arm. Placing a sample in the instrument reduces the magnitude of the electric field of the wave from the sample arm and thus reduces the interferogram proportionately. The amplitude of each wave vector component in the background interferogram and in the sample interferogram can be obtained by Fourier analysis of the interferograms. The ratio of the sample amplitude to the background amplitude for any particular wave vector component is, therefore, the amplitude transmittance of the sample.

The ideas expressed above can be easily symbolized by

$$g(\sigma)=p(\sigma)/p_0(\sigma), \tag{5-1}$$

in which $p_0(\sigma)$ is the Fourier transform of the background interferogram $P_0(x)$,

$$p_0(\sigma)=\int_{-\infty}^{+\infty}\exp(i2\pi\sigma x)P_0(x)\,dx,$$

$p(\sigma)$ is the Fourier transform of the sample interferogram $P(x)$, and $g(\sigma)$ is the amplitude transmittance of the sample. The wavenumber functions are all complex ($p_0(\sigma)$ however, is real in ideal situations) so the amplitude transmittance $g(\sigma)$ contains both the magnitude and phase information.

From Eq. (5-1) by Fourier transformation, we find

$$P(x)=G(x)*P_0(x) \tag{5-2}$$

with

$$G(x)=\int_{-\infty}^{+\infty}\exp(-i2\pi\sigma x)g(\sigma)\,d\sigma$$

and * representing the convolution operation. We see that the sample interferogram is the convolution of a function $G(x)$ with the background interferogram. This $G(x)$ function is the impulse response function for the sample, and $g(\sigma)$ is the spectral response function for the sample. The impulse response function is the output function resulting from an input impulse of Dirac δ-function shape. These response functions are ordinarily considered to be functions of time t and frequency ν, but the substitutions $x=ct$ and $\sigma=\nu/c$, with c the speed of light, give us the optical path difference and the wavenumber functions without difficulty. This allows us to talk, inaccurately of course, as though x were "time" in a new scale and σ a corresponding "frequency."

In order to obtain a physical picture for Eq. (5-2) and, therefore, also for Eq. (5-1), without mathematical involvement, consider an electric wave consisting of Dirac δ-shaped impulses separated by random time intervals and having random signs. Figure 5-3 shows a diagram of such a wave shape, which

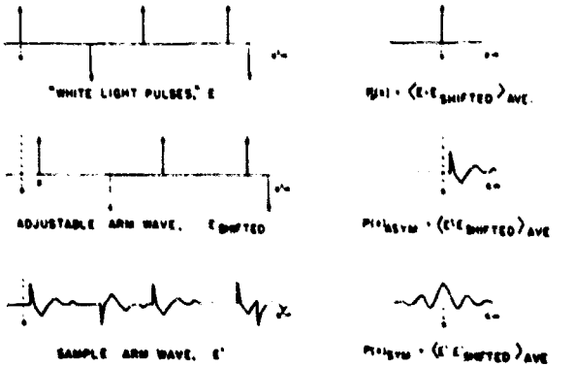


Figure 5-3. Idealized Wave Form of "White Light Pulses" and an Idealized Sample Response to Illustrate the Formation of the Background Interferogram $P_0(x)$, the Asymmetric Mode Interferogram $P_{a.m.}(x)$, and the Symmetric Mode Interferogram $P_{s.m.}(x)$ in a Michelson Interferometer

we will consider to be the wave from the source in our spectrometer. The spectral power of this source wave is uniform over all frequencies, thus the wave is a representation of white light. The splitting, shifting, delaying, and recombining of such a wave in the interferometer would produce an average detector power $P_0(x)$ that has the shape of a Dirac δ-function. With a particular normalization to average over long times x', this $P_0(x)$ interferogram is $\delta(x)$.

The introduction of a sample changes the positive pulses to wavelets of the form $G(x')$ in the sample wave. The sample interferogram will be that portion of the detector power averaged over a long time x' that arises from the product of the wave $\delta(x')$ in the adjustable arm and the wave $G(x')$ in the sample arm. For a delay x in the adjustable arm this interferogram is

$$P(x) = \int_{-\infty}^{+\infty} G(x')\ \delta(x'-x)\ dx'$$

$$= G(x)*\delta(-x)$$

$$= G(x).$$

These waves and interferograms are schematically illustrated in Figure 5-3. We see that the sample interferogram would be the impulse response function if the source wave were "white light pulses," and in this limiting situation Eq. (5-2) is obviously satisfied.

With this special source wave, it is easy to understand the relationship between the symmetric mode interferogram and the asymmetric mode interferogram. The symmetric mode interferogram can be obtained by placing a sample in the source beam or, equivalently, insofar as the detector signal is concerned, by placing identical samples in each of the adjustable and sample arms. The shifted wave in the adjustable arm becomes $G(x'-x)$. The averaged detector signal is

$$\int_{-\infty}^{+\infty} G(x')G(x'-x)\ dx' = G(x)*G(-x),$$

and this is the symmetric interferogram for our special source. The Fourier transform of this symmetric mode interferogram is $g(\sigma)g^*(\sigma) = |g(\sigma)|^2$, with $g^*(\sigma)$ the complex conjugate of $g(\sigma)$, and this is the spectral power transmission function.

The spectral filtering of the detector's nonuniform sensitivity, together with the filtering in the instrument and the fact that the source does not produce "white light pulses," makes the $P_0(x)$ background interferogram have a shape that is not a Dirac δ-function. If the experimental filtering allows a wide band of frequencies through the instrument, $P_0(x)$ is closely packed and peaked at $x=0$ and the interferogram $P(x)$ has many features that are discernable as features of $G(x)$. Regardless of the form of $P_0(x)$, however, Eqs. (5-1) and (5-2) allow the computation of the response functions from the experimental measurements for all frequency components for which $p_0(\sigma)$ is not zero. Further discussion of Eqs. (5-1) and (5-2) are given by Bell[10] and a terse Appendix to this report contains a note about these relations.

The far infrared measurements we made were extended to frequencies that were low enough for us to confidently supply the missing very low and zero frequency characteristics and thereby reconstruct $G(x)$ as the inverse Fourier transform of the $g(\sigma)$ of our measurements. The missing high frequencies in our measurements give us a $G(x)$ that corresponds to a testing impulse with a finite width and missing high frequencies.

Figures 5-4 and 5-5 show examples of such impulse response functions and the test impulse signal. These response functions are obviously causal—no output response preceding the input exciting signal—and thus the dispersion relations (Kramers-Kronig relations)[17,18] must be satisfied by the real and imaginary components of the spectral response functions. Figures 5-4 and 5-5 are presented to show that such functions exist, that they have physical meaning, and because they simplify the description of the behavior of the samples and the measuring and computational techniques.

From a descriptive point of view, $G(x)$ and $G(x)*G(-x)$ are the idealized asymmetric mode and symmetric mode interferograms without the spectral filtering of the measuring equipment. The impulse response functions are real causal functions and are therefore sometimes easier to manipulate mentally

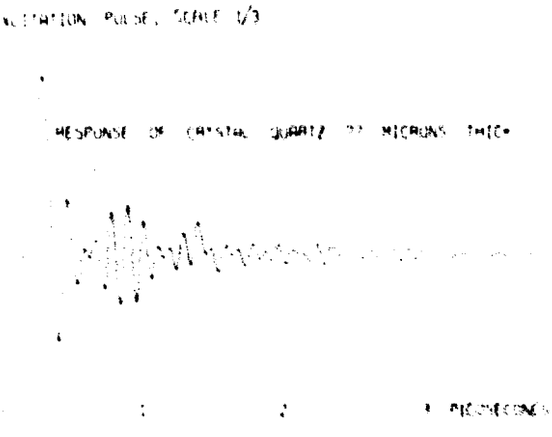


Figure 5-4. Impulse Response Function for the Ordinary Wave Transmitted Through a Lamella of Crystal Quartz. The incident excitation wave pulse is shown reduced in size. A nominal index of refraction is indicated by the delay of the first transmission wave peak

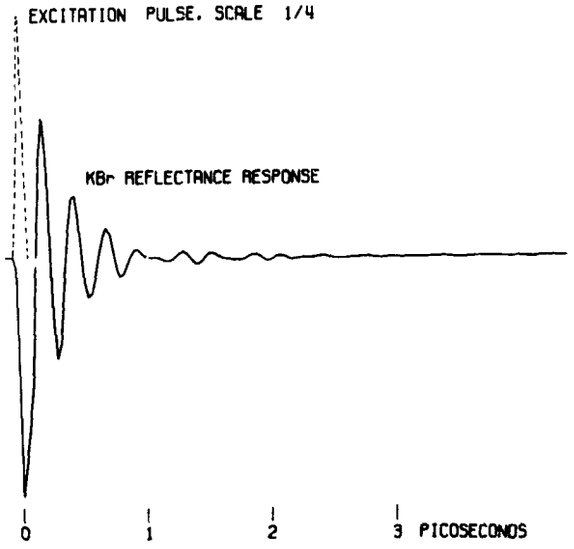


Figure 5-5. Impulse Response Function for the Reflectance From a Potassium Bromide Crystal Surface. The incident excitation wave pulse is shown reduced in size. A nominal index of refraction is indicated by the size of the response during the excitation

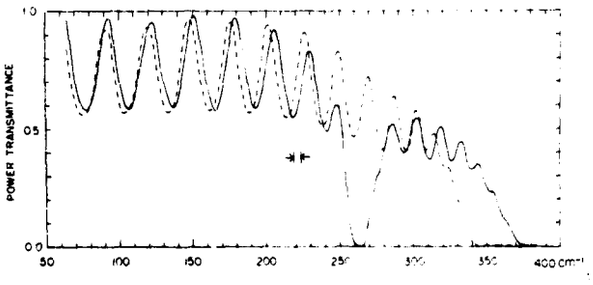


Figure 5-6. The Power Transmittance of the 0.0766 mm Y-cut Quartz Lamella. The ordinary wave is indicated by the solid curve and the extraordinary wave is indicated by the dashed curve.

than the fully equivalent, but complex, spectral response functions. The limitation of the measured interferogram functions to a finite length, for example, has an easily visualized effect upon the measured impulse response function and, therefore, an easily understood effect upon the spectral response function.

5-3 A TRANSMISSION MEASUREMENT OF THE OPTICAL CONSTANTS OF A SOLID

Many of the properties of amplitude spectroscopy are found in the measurement of a plane, parallel faced, solid plate sample. The measurement of such a lamella of crystal quartz by Russell and Bell[12] will be used here as an illustration of the technique. (The full details of the measurement are explained in the Russell and Bell article.) This example will also serve for a comparison of amplitude spectroscopy, in which the measuring instrument is responsive to $G(x)$ and/or $g(\sigma)$, with power spectroscopy, in which the measuring instrument is responsive to $G(x)*G(-x)$ and/or $|g(\sigma)|^2$.

The impulse response function shown in Figure 5-4 was measured for a Y-cut sample of crystal quartz, $76.6 \pm 0.3\mu$ thick, with a polarizer oriented so that the ordinary wave produced both the sample and the background interferograms. The independent variable has been relabeled in the figure as time in picoseconds, rather than optical path difference, to emphasize the physical origin of the impulse response function. The power transmittance of this sample is shown in Figure 5-6, as obtained by squaring the magnitude of the amplitude transmittance. This is equivalent to the information that would have been obtained from a power spectroscopic measurement. The power spectrum for the extraordinary wave is also shown in the figure. The channeled spectra are evident in both spectra.

The index of refraction may be obtained from the position, the spacing, and the magnitude of the channeled spectrum[19,20] but care must be exercised[20,21] and the sample must be reasonably transparent. At the channeled spectrum transmission maxima, for example, the optical thickness of the sample is an integral number, the order number, of half wavelengths. The accuracy of the index of refraction determination at the frequency of a maximum depends upon the precision of the location of the maximum; this precision is affected by the resolution and by the superimposed absorption structure in the spectrum. Amplitude spectroscopy, on the other hand, gives the transmission spectral response function $T(\sigma)$ in magnitude and in phase at each frequency

$$g(\sigma) = |g(\sigma)| \exp i\phi_g(\sigma) = |T(\sigma)| \exp i\phi_T(\sigma).$$

The effects of resolution and absorption in the index of refraction determination, which is obtained mainly from the phase, are very much reduced in comparison to power spectroscopy.

The index of refraction n and the extinction coefficient k of the complex index $n+ik$ are approximately related to the amplitude transmittance of the lamella by

$$|T| \exp i\phi_T = (1-r^2) \exp(-2\pi\sigma kb) \exp i(2\pi\sigma(n-1)b) \quad (5\text{-}3)$$

with r the complex amplitude external reflectance of the bulk material. This expression is valid insofar as

only the first partial wave through the sample is important, and the multiple internally reflected partial waves may be neglected. The first partial wave produces a "signature" in the interferogram at an optical path displacement $x=(n-1)b$. The second partial wave arrives later at an optical path displacement $x=(3n-1)b$, and the higher partial waves arrive even later.

A low resolution amplitude spectroscopic measurement, which is made with a short interferogram that does not include the $(3n-1)b$ second partial wave "signature," can be analyzed exactly with the above approximation—no channel spectrum will appear at this resolution. With r^2 small compared to unity, the n and k can be separately determined from

$$n=1+\phi_T/2\pi\sigma b$$

and

$$k=(1/2\pi\sigma b)\ln(1/|T|).$$

This approximation can be improved by using n and the approximated value of k to compute the reflectance by

$$r=(1-(n+ik))/(1+(n+ik))$$

which, together with the measured amplitude transmittance, gives

$$(1-r^2)^{-1}|T|\exp i\phi_T=|T'|\exp i\phi_{T'}$$
$$=\exp(-2\pi\sigma k'b)$$
$$\times\exp i(2\pi\sigma(n'-1)b)$$

so that the improved values n' and k' can be obtained. Iteration of this process quickly produces optical constant values with an accuracy appropriate to the experimentally measured data. The iteration process can start with any reasonable approximate values of the optical constants; the accurate values for a neighboring frequency are an appropriate choice when the constants are to be calculated for a whole range of frequencies.

The possibility of measuring the index of refraction at very low spectral resolution by amplitude spectroscopy is an experimental convenience that is not available in power spectroscopy. The low resolution, obviously, measures some mean value of the index of refraction over a frequency interval determined by the spectral slit function. The resolution necessary to measure a particular sample depends upon the amount of structure in the spectrum of the sample and the accuracy required of the measurement. Smaller spectral slit widths require longer interferograms and the inclusion of more partial waves in the measurement and analysis. The expression for the transmission of the lamella given in Eq. (5-3) can be extended[12] to include the number of partial waves that are measured in the interferogram; the effects of apodization may also be included. Calculation with a simple extension of the iteration procedure indicated above, which is a trivial task for any general purpose digital computer that can calculate the Fourier transformation of the original interferograms, gives the optical constants of the sample.

Figure 5-7 shows the optical constants n and $\alpha=4\pi\sigma k$ for quartz, as measured by Russell by this amplitude spectroscopic method at the spectral resolution indicated in Figure 5-6. Since the computation of ϕ_T from the interferograms gives only the principal value, it is necessary to establish the number (the order number) of unobserved complete cycles. In the example of the ordinary wave in quartz, it is difficult to follow the changing phase through the absorption band at 264 cm^{-1}, because the signal is weak in the band and the phase shift is large in passing through the band. Since the dispersion relations relate the real and imaginary parts of the impulse response function, T relates with ϕ_T and n with α; consequently, meas-

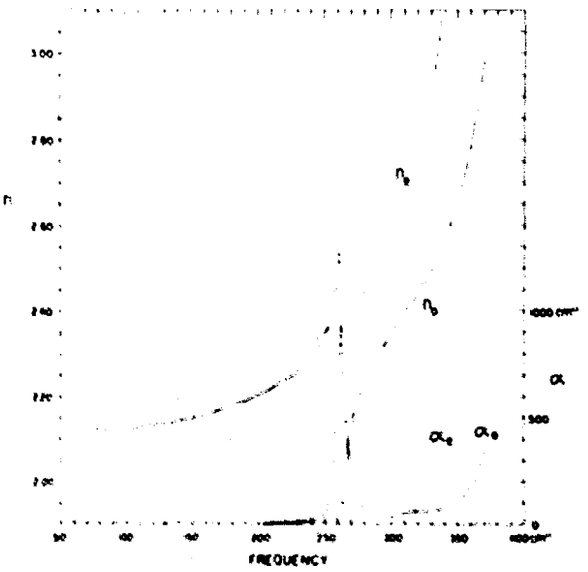


Figure 5-7. The Ordinary-ray (solid curves) and Extraordinary-ray (dashed curves) Refractive Indices and Absorption Coefficients of Quartz as Determined From the 0.0766 mm Lamella by Amplitude Spectroscopy

urement of the absorption can be used to establish the change in order number across the band. Since the phase shift across the quartz band was verified by this procedure, an accurate value of the shift was known and an accurate value of the band absorption strength could be determined. It is evident that the use of α to help in establishing n, as in this quartz example, can proceed only in those special cases where α is measurable with sufficient accuracy. For very large phase shifts the method fails because the corresponding bands are too opaque for measurement.

In every phase measurement it is necessary to establish the order number in order to know the number of complete cycles that must be added to the measured phase. This can usually be accomplished by using an approximate index of refraction, a measurement of samples of different thicknesses, or some such additional information.

The ambiguity in the phase information is easily removed in the very long wave region of the spectrum where an extrapolation to zero frequency, and thus to a zero phase, is possible. As an example of the precision of such an amplitude spectroscopic measurement of an index of refraction, Figure 5-8 shows the

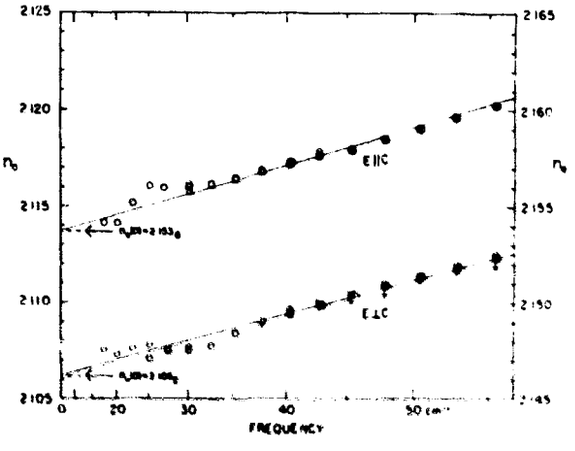


Figure 5-8. The Ordinary- and Extraordinary-ray Refractive Indices of Quartz as Measured by Amplitude Spectroscopy. The extrapolated, zero frequency values have an experimental uncertainty of ±0.001, which is much larger than is apparent from the consistency of the data

long wave result for crystal quartz. The abscissa scale was chosen as the square of the frequency so that the index values would extrapolate along a straight line to zero frequency. The circles are experimental points obtained for a sample of quartz 4.78?±0.0005 mm thick. It is seen that the scatter of points at the higher frequencies in the figure is not much greater than the uncertainty in $n-1$ arising from the uncertainty in the thickness of the sample. A small correction was made in these index values for the finite solid angle of the radiation through the sample—the optical path length through the sample was not the same for all rays.

The precision of the index of refraction measurement depends upon the precision of the determination of both the sample's thickness and the phase of the spectral response function. The new problem in amplitude spectroscopy is to evaluate the effect of the detector system noise on the phase precision. As in the case of ordinary spectra, the noise can be judged from the results of many measurements at the same frequency or at closely neighboring frequencies in a phase spectrum. It is also possible to predict the phase noise, just as it is possible to predict the noise in the magnitude of the spectral response function, from the known noise in the interferogram and some knowledge of the size of the magnitude spectrum. With the stepwise measurement of the interferogram, as used by Russell, the noise value at any interferogram point is independent of the noise value at any other point. It is then easy to estimate the noise in the spectrum. The noise in amplitude spectroscopy is a complex value having both magnitude and phase that adds to the complex spectral signal. The uncertainty in the phase in radians is just the noise to signal ratio, if this uncertainty is a small fraction of a radian.[10] Thus, with a little knowledge, it is possible to predict the goodness of the phase data that would be obtained with various experimental conditions: number of interferogram points, spacing of the interferogram points, time for measurement, etc., just as one can for the magnitude or for power spectroscopy.

An important feature of amplitude spectroscopy, used with the radiation passing throughout the sample only once, is the ability to measure small transmittance values. The power transmittance of the quartz sample shown in Figure 5-6, for example, was 0.0013 at 264.6 cm^{-1}. This value would be extremely difficult to measure by power spectroscopy. The magnitude of the amplitude transmittance, however, was 0.036, a measurable value. This important feature of amplitude spectroscopy is illustrated by the data in Table 5-1 for two hypothetical samples: a "trans-

Table 5-1. Comparison of the Signal and Noise Output of Two Samples by Power Spectroscopy and by Amplitude Spectroscopy

Sample	Power Spectroscopy	Amplitude Spectroscopy
"Transparent" Sample:		
background signal	100±1	100±1
sample signal	100±1	100±1
$\|T\|$		1.00±0.02
$\|T\|^2$	1.00±0.02	1.00±0.04
"Opaque" Sample:		
background signal	100±1	100±1
sample signal	1±1	10±1
$\|T\|$		0.10±0.01
$\|T\|^2$	0.01±0.01	0.010±0.002

parent" 100 percent power transmitter, and an "opaque" 1 percent power transmitter. Both are measured with an assumed incident beam background of 100 units strength in the same interferometer with an assumed ±1 unit of noise. The measurements with the samples also give ±1 unit of noise and 100 units of output with the "transparent" sample, 1 unit of output with the "opaque" sample by power spectroscopy, and 10 units of output with the "opaque" sample by amplitude spectroscopy.

The table indicates the signal to noise superiority for power spectroscopy of the "transparent" sample and the signal to noise superiority for amplitude spectroscopy of "opaque" samples. It should be emphasized here that far infrared Fourier transform spectroscopy with a Michelson interferometer has two great advantages over normal spectrometers: the well known "Fellgett advantage" in signal to noise, and a very greatly reduced stray radiation advantage.* This second advantage helps contribute to the ability of Fourier transform spectroscopy to cope with opaque samples.

The ability of amplitude spectroscopy to measure power transmittance is completely lost if the sample lamella does not have high-quality plane-parallel faces. A "pseudocoherence" effect[16] exists in amplitude spectroscopy that does not exist in power spectroscopy. This effect reduces the magnitude of the amplitude transmittance for samples having low geometrical quality. One way to visualize this pseudocoherence effect is to consider dividing the sample area into smaller areal sections and measuring the interferograms of each small section separately. The interferogram $P(x)$ for the full area sample is the sum of the interferograms $\Delta P(x)$ for all of the sections. Thus the Fourier transform $p(\sigma)$ is the sum of the $\Delta p(\sigma)$ elements. Since the $\Delta p(\sigma)$ elements are complex, the magnitude of $p(\sigma)$ is not the sum of the magnitudes of the $\Delta p(\sigma)$ elements. The resulting $G(x)$ and $g(\sigma)$ are seen to be an average of the real $\Delta G(x)$ elements and an average of the complex $\Delta g(\sigma)$ elements.

A power spectroscopic measurement for the corresponding experiment would yield symmetric interferograms $\Delta P(x)$ for the elemental sections and real positive $\Delta p(\sigma)$ spectral densities. The resulting values $G(x)*G(-x)$ and $g(\sigma)g^*(\sigma)$ are an average of real symmetric $\Delta(G(x)*G(-x))$ elements and an average of real positive-valued $\Delta(g(\sigma)g^*(\sigma))$ elements.

The power transmittance obtained from an amplitude spectroscopic measurement is the square of the magnitude of the average amplitude transmittance; this is quite different from the power spectroscopic measurement of the average of the squares of the magnitudes of the transmittances. The distinction is illustrated by the result[16] shown in Figure 5-9 in which the crosses show a power spectroscopic measurement and the curve an amplitude spectroscopic measurement of the same sample of polyethylene containing carbon black. The sample had a thickness varying between 1.53 and 1.61 mm, the 0.08 mm

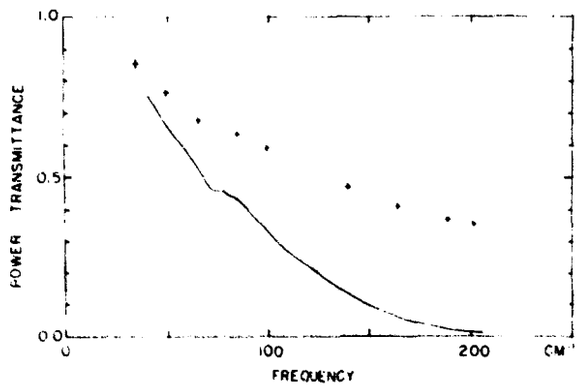


Figure 5-9. The Crosses Show the Power Transmittance of a Sample of "Black" Polyethylene Measured With a Spectrometer. The solid curve shows the measurement results for the same sample by amplitude spectroscopy. The amplitude spectroscopy results are spoiled by the thickness variation, 0.153 to 0.161 mm, of the sample

variation being only a 5 percent variation in the thickness. Such a sample is adequate for power spectroscopic measurements, but very bad for the amplitude spectroscopic measurement. At 200 cm^{-1} this variation in thickness is about 1 wavelength of insertion optical path, $(n-1)b$, and, therefore, the elemental $\Delta g(\sigma)$ values have phases that are distributed over a full cycle. The sum of these $\Delta g(\sigma)$ values is nearly zero; thus $g(\sigma)$ and $|T|^2$ are nearly zero.

The 200 cm^{-1} radiation passes through the sample but does not produce any contribution to the asymmetric mode interferogram because the phases of the 200 cm^{-1} components of the elemental $\Delta P(x)$ interferograms are distributed over a full cycle and, therefore, destructively interfere. An even more spectacular example of the pseudocoherence of different polarization components in crystal quartz has been given by Bell[16] and by Russell.[12]

5-4 A REFLECTION MEASUREMENT OF THE OPTICAL CONSTANTS OF A SOLID

Materials that are so strongly absorbing that they cannot be measured in transmission may be measured in reflection. There is no channeled spectrum, of course, to help in evaluating the index of refraction. If, however, in addition to the power reflectivity $|r|^2$ the phase of the amplitude reflectivity ϕ_r were known, then the optical constants could be calculated from the relations

$$n = (1-|r|^2)/(1+|r|^2+2|r|\cos\phi_r) \qquad (5\text{-}4)$$

and

$$k = (2|r|\sin\phi_r)/(1+|r|^2+2|r|\cos\phi_r). \qquad (5\text{-}5)$$

* The asymmetric mode has an additional advantage of producing no interferogram from the thermal emission from the sample

A well established procedure in power spectroscopy is to derive the phase of the reflectivity from the power reflectivity spectrum. Thus the reflectivity phase is obtained by use of the dispersion relations[17,18] as

$$\phi_r(\sigma) = (\sigma/\pi)\int_0^\infty d\sigma'(\ln|r(\sigma)|^2 - \ln|r(\sigma')|^2)/(\sigma'^2-\sigma^2),$$

where $|r(\sigma)|^2$ is the power reflectivity at the frequency σ and the Cauchy principal value of the integral is used to properly include the singularity at $\sigma'=\sigma$. The tribulations of such calculations have been well described[22,23]. The inaccuracy of the measurement of the low reflectivity region on the high frequency side of strong reflections produces errors in the large values of $\ln|r|^2$ and cause great difficulty with the technique. The usual lack of reflectivity data throughout a large range of frequencies—zero to infinity required in the formula—is a smaller problem, but is a nuisance, nevertheless.

With amplitude spectroscopy, both the magnitude and the phase of the reflectivity are directly determined at each measured frequency. Replacing a mirror in the sample path arm with the reflecting surface of a thick sample allows the magnitude and phase of the sample reflectivity to be compared with those of the mirror (π radians phase). The spectral response function is $g(\sigma) = |r(\sigma)| \exp i(\phi_r(\sigma) - \pi)$. The measured phase, of course, also depends upon the position of the sample's reflecting surface, which is supposed to exactly replace the mirror surface position. If the phase of the sample's reflectivity is known in a high frequency region, then a comparison of the measured phase with the known phase can serve to determine the position error; such an error is easily removed from the phase data (or the impulse response function) by an appropriate correction.

The results obtained by Johnson[24] on the optical constants of crystal potassium bromide will be used to illustrate reflectivity measurements by amplitude spectroscopy. The impulse response function shown in Figure 5-5 is for potassium bromide reflectance. The "ringing" of the crystal at its eigenfrequency is clearly evident. The spectral power reflectivity $|r(\sigma)|^2$ of potassium bromide, which could have been measured by power spectroscopy, and the spectral phase of the reflectance $\phi_r(\sigma)$, which cannot be measured by power spectroscopy, are shown in Figure 5-10. The error bars on the curves indicate the scatter of the measured values in the various frequency regions. The amount of structure in the phase spectrum in the region of low reflectivity, where the dispersion relations are difficult to use, should be noted. The fact that the phase can be measured in the regions of very weakly reflected radiation should also be noted. This ability of amplitude spectroscopy to measure weak signals is also indicated in Figure 5-11, which gives the power reflectance measured in this weakly reflecting region, and was explained in connection with Table 5-1. The measured amplitude reflectivity does agree with the approximate value $-i\lambda/2$ at the minimum.

The optical constants of crystal potassium bromide derived from the reflection measurement are shown in Figure 5-12. The index of refraction and the extinction coefficient were calculated from the measured amplitude reflectivity according to Eqs. (5-4) and (5-5). The physical significance of the results are discussed in the paper by Johnson.[24]

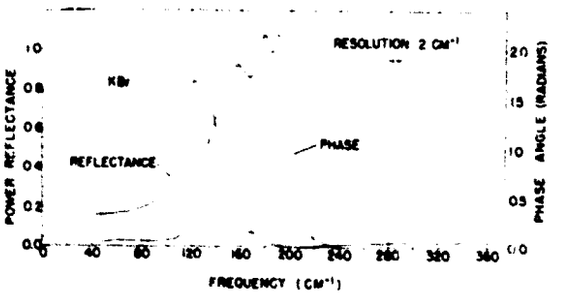


Figure 5-10. The Power Reflectance $|r(\sigma)|^2$, and the Phase of the Reflectance $\phi_r(\sigma) - \pi$, for Potassium Bromide as Measured by Amplitude Spectroscopy. The error bars on the curves show the size of the variation in the measured values for several different sample measurements

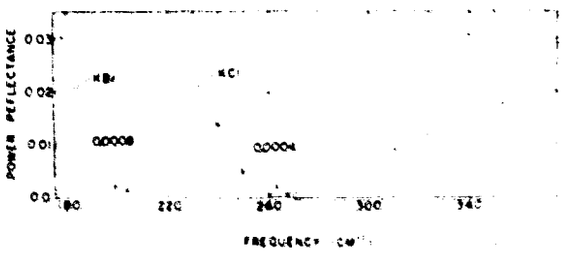


Figure 5-11. The Power Reflectance $|r(\sigma)|^2$ for Potassium Chloride and for Potassium Bromide in a Low Reflectance Region to Show the Ability of Amplitude Spectroscopy to Measure Small Power Reflectances. The error bars on the curves show the variations in the measured values for several sample measurements

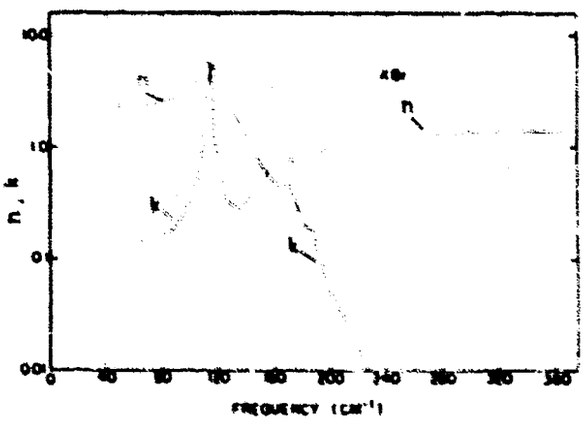


Figure 5-12. The Index of Refraction $n(\sigma)$ and Extinction Coefficient $k(\sigma)$ of Potassium Bromide Determined by Amplitude Spectroscopy

5-5 AMPLITUDE SPECTROSCOPIC MEASUREMENTS ON GASES AND LIQUIDS

The usefulness of amplitude spectroscopy in the measurement of the strengths of far infrared absorption lines in gases has been demonstrated by Chamberlain[6] and by Sanderson[14,15]. A paper that will be presented by Sanderson later in this Conference (Chapter 14) will indicate the high degree of development of this technique. It is appropriate here, however, to look at an example of Sanderson's earlier work shown in Figure 5-13. The figure shows the index of refraction of hydrogen chloride gas that was

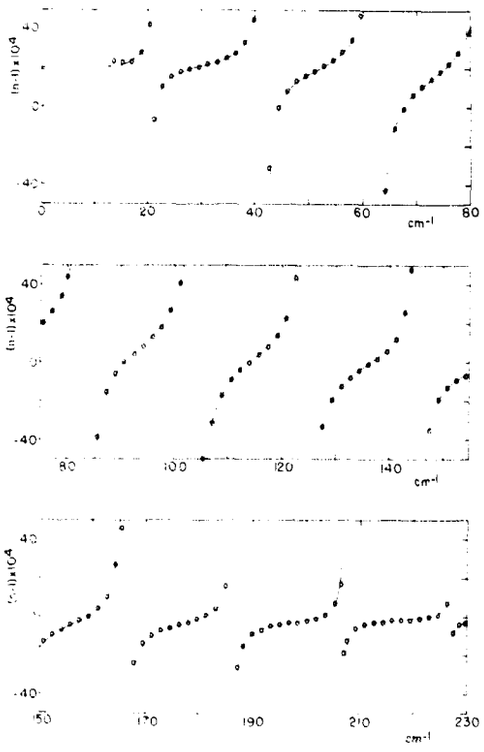


Figure 5-13. The Index of Refraction per Atmosphere of Hydrogen Chloride Gas as Measured by Sanderson[14] (1967) by Amplitude Spectroscopy. The circles are experimentally determined values; the solid line is calculated from the line strengths as determined from the experiment

obtained from the phase spectrum, $(n-1)=\phi/(2\pi\sigma b)$, measured with the interferometer described earlier. From the experimental index values, the strengths of each of the lines (at 20 cm^{-1}, 40 cm^{-1}, 60 cm^{-1}, etc.) was determined. These line strengths were used to calculate the index values shown by the solid line in the figure. The important feature of amplitude spectroscopy for this work is the fact that the line strength can be measured by the index of refraction changes in the neighborhood of the line. Because the gas absorption lines are narrower than typical spectrometer slit widths, power spectroscopic measurements are made with amounts of gas such that the lines are opaque at their centers and, thus, information about their strengths is lost. The direct measurement of the index of refraction and the line strengths is valuable additional information.

Amplitude spectroscopic measurements of liquids have been made by Chamberlain et al[7,9]. It is difficult to construct a cell for liquids that can be placed in a vacuum far infrared instrument and that has the necessary precision in its windows and spacing so that the "pseudocoherence" effect will not cause troubles. Chamberlain has solved this problem by measuring a liquid layer supported on a level mirror; the mirror serves as the end mirror in the sample arm. The radiation incident from above on this liquid layer is transmitted through the liquid and reflected by the mirror. The wave from the sample arm to the detector contains a partial wave that is reflected from the top liquid surface, a partial wave that is reflected once from the end mirror, and partial waves that are reflected several times from the end mirror and internally reflected from the top liquid surface. These several partial waves can be seen in Figure 5-14, which is an interferogram recorded and described by Chamberlain et al[7].

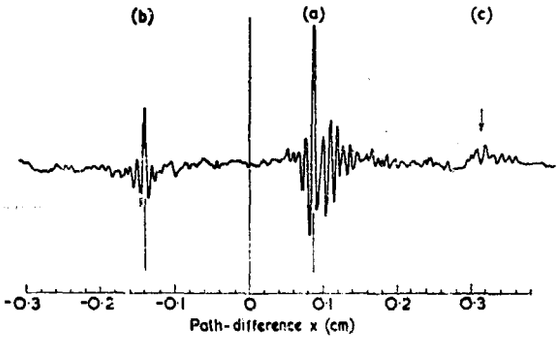


Figure 5-14. An Asymmetric Mode Interferogram, Measured by Chamberlain et al, (1967) for a 0.68 mm Thickness of Liquid Tetrabromoethane on a Mirror. The signature (b) is due to the partial wave reflected from the top surface of the liquid; the signature (a) is due to the partial wave transmitted twice through the liquid layer; the signature (c) is due to the partial wave that is reflected internally once from the top liquid surface, reflected twice by the mirror, and transmitted four times through the liquid layer. Reprinted with permission from Chamberlain et al., (1967) *Spectrochimica Acta*, **23A**: 2256, Pergamon Press

To simplify the analysis of the transmission through the liquid sample, Chamberlain removed the top surface, first reflection signature from the interferogram (by smoothing over the signature region) before calculating the Fourier transform. The remaining interferogram, of course, is not the same as would be obtained in transmission with a double-thickness liquid layer, because of the mirror precisely

centered in this double-thickness liquid path. The partial wave analysis, however, allows the calculation of the optical constants to the desired accuracy.

If a separate measurement of the reflected partial wave alone is measured, by canting the end mirror to remove the transmitted wave, for example, then the optical constants could be obtained from the reflected partial wave alone. It would also be possible, then, to remove the reflected partial wave signature in all details from the transmission interferogram by subtraction of the one interferogram from the other.

Such possibilities add a new freedom to amplitude spectroscopy that is not so readily available in power spectroscopy. It should be noted, however, that a preprocessing of the interferogram can also be an aid in the analysis of a power spectroscopic measurement. Randall[19], for example, increased the channeled spectrum signal relative to the noise by removing portions of the interferogram that did not contribute to the channeled spectrum structure. The preprocessing of the symmetric interferogram improved the channeled spectrum measurement and the determination of the index of refraction of the sample.

5-6 CONCLUSION

Amplitude spectroscopy is a useful technique for measuring the optical constants of materials that can be obtained in samples of sufficiently precise geometrical form.* The measurement at each frequency of a magnitude and a phase allows the direct computation of a pair of optical constants at each frequency. Because the optical constants of a material primarily describe the behavior of a wave amplitude in its passage through the material, amplitude spectroscopy generally gives results more simply related to the optical constants than does power spectroscopy. The measurement of a wide band of frequencies from a source results in a close-packed background interferogram and the development of separated signatures for the partial waves from the sample that can be useful in analysis. The combination of a general purpose digital computer with a modest asymmetric mode interferometer can measure optical constants with a useful accuracy.

The large signal to noise improvement of Fourier transform spectroscopy in the infrared, the spectral purity of the Michelson interferometer measured interferogram, and the large value of the amplitude transmittance or reflectance compared with the power transmittance or reflectance make infrared amplitude spectroscopy with an asymmetric Michelson interferometer especially advantageous for measuring strongly absorbing materials.

* On the other hand, the technique might be used to obtain a measure of the geometrical irregularities—surface roughness, for example.

Appendix

The response of the interferometer with different samples in the two arms may be formulated in the following simple manner. Let the electric field source wave into the sample arm be $V_1(x)$ and the wave into the adjustable arm be $V_2(x)$. Here x is a "time" or the equivalent optical path distance. The background interferogram function will be the average power $P_0(x)$ given by, except for uninteresting constants, $P_0(x) = V_1(x) * V_2(-x)$. The addition of samples in each of the arms will change the output wave from the arms to

$$V_1'(x) = G_1(x) * V_1(x) \quad \text{and} \quad V_2'(x) = G_2(x) * V_2(x),$$

with $G_1(x)$ and $G_2(x)$ the sample impulse response functions. The interferogram function with the samples in place will be

$$\begin{aligned} P(x) &= V_1'(x) * V_2'(-x) \\ &= (G_1(x) * V_1(x)) * (G_2(-x) * V_2(-x)) \\ &= (G_1(x) * G_2(-x)) * (V_1(x) * V_2(-x)) \\ &= (G_1(x) * G_2(-x)) * P_0(x). \end{aligned}$$

The corresponding Fourier-transformed relation is

$$\begin{aligned} p(\sigma) &= v_1'(\sigma) v_2'^*(\sigma) \\ &= g_1(\sigma) g_2^*(\sigma) p_0(\sigma), \end{aligned}$$

in which $v_1'(\sigma)$ is the Fourier transform of $V_1'(x)$, etc. If the adjustable arm contains a sample identical to

that in the sample arm, then the result is the expected result for symmetrical sample operation, power spectroscopy, except that the background has not been presumed to be symmetric. If the adjustable arm contains no sample, then $G_2(x)=\delta(x)$ and $g_2(\sigma)=1$, with the obvious simplification to the asymmetric case (amplitude spectroscopy) and there is no insistence upon a symmetric background function.

It is interesting to note that

$$G(x)=\delta(x)-2\pi b\, d\chi(x)/dx$$

for a sufficiently thin specimen, where $\chi(x)$ is the susceptibility of the material. In this approximation $G(x)*G(-x)$ is just twice the symmetrical portion of $G(x)$. For thick samples this simplicity is lost.

References

1. Connes, J., (1961) *Revue Opt. Theor. Instrum.* **40**: 45, 116, 161, 231.

2. Vanasse, G. A., and Sakai, H., (1967) *Progress in Optics VI* E. Wolf, ed., North Holland Pub., Amsterdam.

3. Chamberlain, J. E., Gibbs, J. E., and Gebbie, H. A., (1963) *Nature* **198**: 874.

4. Chamberlain, J. E., (1965) *Infrared Physics* **5**: 175.

5. Chamberlain, J. E., and Gebbie, H. A., (1966) *Applied Optics* **5**: 393.

6. Chamberlain, J. E., and Gebbie, H. A., (1965) *Nature* **208**: 480.

7. Chamberlain, J. E., Costley, A. E., and Gebbie, H.A., (1967) *Spectrochimica Acta* **23A**: 2255.

8. Chamberlain, J. E., (1967) *J. Quant. Spect. Radiative Transf.* **7**: 151.

9. Chamberlain, J. E., Gebbie, H. A., Pardoe, G. W. F., and Davies, M., (1968) *Chem. Phys. Let.* **1**: 523.

10. Bell, E. E., (1966) *Infrared Physics* **6**: 57.

11. Russell, E. E., and Bell, E. E., (1966) *Infrared Physics* **6**: 75.

12. Russell, E. E., and Bell, E. E., (1967) *J. Opt. Soc. Am.* **57**: 341.

13. Russell, E. E., and Bell, E. E., (1967) *J. Opt. Soc. Am.* **57**: 543.

14. Sanderson, R. B., (1967) *Appl. Optics* **6**: 1527.

15. Robinette, W. H., and Sanderson, R. B., (1969) *Applied Optics* **8**: 711.

16. Bell, E. E., (1967) *J. Physique, Colloque C-2*, **28**: C2-18.

17. Toll, J. S., (1956) *Phys. Rev.* **104**: 1760.

18. Stern, F., (1963) *Solid State Physics* **15**: 299, Seitz and Turnbull ed. Academic Press, New York.

19. Randall, C. M., and Rawcliffe, R. D., (1967) *Applied Optics* **6**: 1889.

20. Bell, E. E., (1967) *Handbuch der Physik, Bd XXV/2a* Genzel ed., Springer Verlag, Heidelberg.

21. Moss, T. S., (1957) *Proc. Phys. Soc. (London)* **B70**: 778.

22. Spitzer, W. G., and Kleinman, D. A., (1961) *Phys. Rev.* **121**: 1324.

23. Wu, C., and Andermann, G., (1968) *J. Opt. Soc. Am.* **58**: 519.

24. Johnson, K. W., (1969) *Phys. Rev.* **187**: 1044.

Contents

6. Computing Problems in Fourier Spectroscopy

Janine Connes
Centre Inter-Disciplina Regional de Calcul
Electronique, C.N.R.S.
91 Campus D'Orsay
France

Abstract

The problems encountered in data handling and processing of interferograms are discussed and solutions to these problems are presented. Computing techniques and programs are described and the "decimation in time" approach explained. A comparison of the computer times required for the FFT (fast Fourier transform) for various types of computers is tabulated. Finally, a description of some operations performed on computed spectra is given.

6–1 INTRODUCTION

The superiority of Fourier spectroscopy over grating spectroscopy has been shown in widely different circumstances. We shall merely mention two extreme cases, the very low resolution spectra taken by the Nimbus satellite containing merely a few hundred spectral elements, and very high resolution laboratory spectroscopy as practiced at Aimé Cotton, with 10^6 spectral elements.

Problems encountered in interferometer building are totally different in both cases. The same is true of computing techniques. However in any case, and with any computer, one can always consider five groups of problems. We shall put them in the order in which they occur during an actual experiment:

(1) Doctoring the actually recorded interferogram to correct for various instrumental errors and make it close to ideal.

(2) Rapid computation of a small spectrum slice from the entire interferogram for testing purposes.

(3) Computation of the entire spectrum from an ideally odd or even interferogram.

(4) Operations on the computed spectrum (apodizat[illegible]ation, averaging of several spectra).

(5) Extrac[illegible]ng the data from the spectra, that is, automatic detection of lines, and measurements of line positions, intensities, profiles, and so forth.

6–2 SAMPLING AND RESOLUTION

Let $B(\sigma)$ be the spectrum under study. The light falls on a two beam interferometer with variable path difference δ; the output light flux can be written as[1]

$$\Phi(\delta)=\int_0^\infty B(\sigma)\,d\sigma+\int_0^\infty B(\sigma)\cos 2\pi\sigma\delta\,d\sigma.$$

The variable part $s(\delta)$, termed the interferogram, is recorded. It is the autocorrelation function of the light vibration,[2]

$$s(\delta) = \int_0^\infty B(\sigma) \cos 2\pi\sigma\delta \, d\sigma.$$

In some cases, in order to cancel light source fluctuations, "internal modulation" is used.[3,4] The signal is then, to a first approximation, the derivative of the previous one and the constant part is automatically eliminated:

$$I(\delta) = \int_0^\infty B(\sigma) \sin 2\pi\sigma\delta \, d\sigma.$$

With an ideally built and adjusted interferometer, $I(\delta)$ is strictly an odd or even function, and the even $B_e(\sigma)$ or odd $B_o(\sigma)$ parts of the spectrum can be reconstructed by performing the inverse operation:

$$B_e(\sigma) = \int_0^\infty s(\delta) \cos 2\pi\sigma\delta \, d\delta$$

$$B_o(\sigma) = \int_0^\infty I(\delta) \sin 2\pi\sigma\delta \, d\delta, \qquad (6\text{-}1)$$

according to whether $I(\delta)$ is even or odd (Figure 6-1).

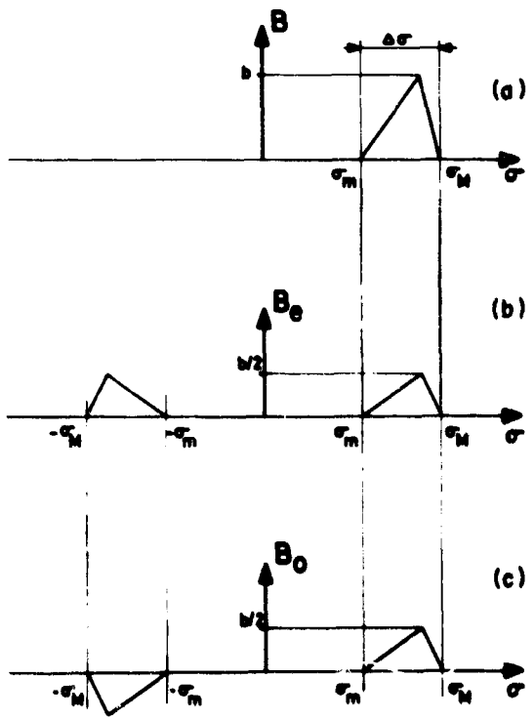


Figure 6-1. (a) Optical Spectrum Under Study; (b) Even Part of the Optical Spectrum Computed Through a Cosine FT of the Even Interferogram; (c) Odd Part of the Optical Spectrum Computed Through a Sine FT of the Odd Interferogram

The techniques actually used for reconstructing the spectrum are in the large majority of cases purely digital, and operate from sampled interferograms. Only a finite number of samples can be used both for the interferogram and the spectrum. We shall show that, provided suitable care is used, it is possible to reconstruct the spectrum corresponding to a given maximum path difference δ_{max} without making any approximation.

We shall consider the case of the sine FT of an odd interferogram. The same results would apply to the cosine FT of an even interferogram.

6–2.1 Sampling Interval

All spectra are limited by source emission, optics transmission, receiver sensitivity, or by filters. Let us call σ_1 and σ_2 the limits of the frequency range, and put $\Delta\sigma = \sigma_2 - \sigma_1$.

According to the sampling theorem all the information about a band limited function is given by a set of discrete samples of the function and interpolation formulas can be found which reconstruct the continuous function.[1,5-8]

Let h be the sampling interval. Instead of computing the integral (6–1) we shall compute the sum

$$B_o'(\sigma) = \sum_{k=0}^{k=\infty} I_{kh} \sin 2\pi\sigma kh, \qquad (6\text{-}2)$$

which can be considered as the integral:

$$B_o'(\sigma) = \int_0^\infty I(\delta) R_h(\delta) \sin 2\pi\sigma\delta \, d\delta, \qquad (6\text{-}3)$$

where $R_h(\delta)$ is a Dirac distribution with a period equal to h (Figure 6-2).

The sine FT of the product between two odd and even functions is the convolution of the sine FT of the odd function by the cosine FT of the even function.

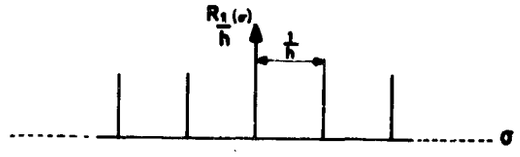


Figure 6-2. Dirac Distribution $R_h(\delta)$ with Periodicity h

The sine FT of $I(\delta)$ is $B_o(\sigma)$, the cosine FT of $R_h(\delta)$ is within a factor h another Dirac distribution $R_{1/h}(\sigma)$ with periodicity $1/h$ (Figure 6-3).

$$B_o'(\sigma) = hB_o(\sigma) * R_{1/h}(\sigma).$$

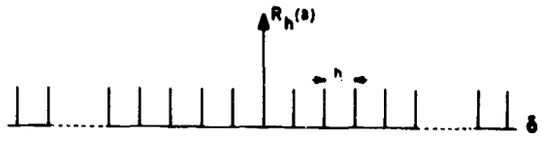


Figure 6-3. Dirac Distribution $R_{1/h}(\sigma)$ with Periodicity $1/h$

This means that the odd spectrum $B_o(\sigma)$ is periodic with periodicity $1/h$.

Figures 6-4 and 6-5 present the sine FT of an odd interferogram and the cosine FT of an even interferogram. They show that within the sampling

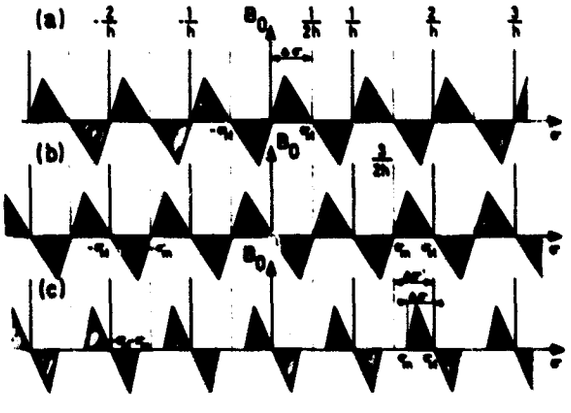


Figure 6-4. Aliasing of the Spectra Computed from an Odd Interferogram.
(a) Optical spectrum in the range 0 to σ_M;
(b) Optical spectrum in the range σ_m to σ_M, with $\sigma_m = K\,\Delta\sigma$, K being an integer;
(c) Optical spectrum in the range σ_m to σ_M, with $\sigma_m \neq K\,\Delta\sigma$

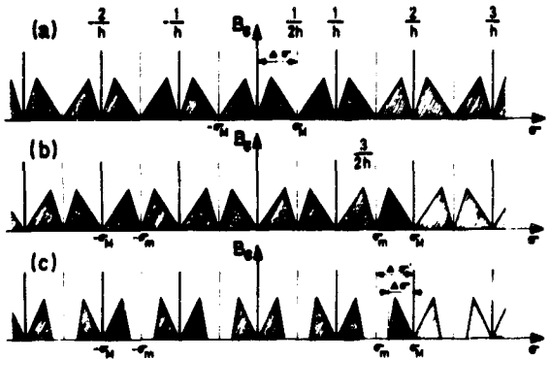


Figure 6-5. Aliasing of the Spectra Computed from a Sampled Even Interferogram.

interval $1/h$ one finds both the wanted spectrum and an "image" or "alias". This is one way of showing that the complex FT of a real function is an Hermitian function (Appendix A). In order to avoid overlapping one must select

$$\text{either } h = \frac{1}{2\,\Delta\sigma} \quad \text{(case } a \text{ and } b\text{)},$$

$$\text{or} \quad h = \frac{1}{2\,\Delta\sigma'} \quad \text{(case } c\text{, with } \Delta\sigma' > \Delta\sigma\text{)}.$$

If this is done, within the interval $\frac{1}{2h}$ which corresponds to the range of interest, one finds the original spectrum without any kind of approximation.

6-2.2 Wavenumber Resolution

The path difference being limited to δ_{max}, only $N = \delta_{max}/h$ interferogram samples will be used for computing $B(\sigma)$ as given by (6-3); then

$$B(\sigma) = \sum_{k=0}^{k=N-1} I_{k\Delta} \sin 2\pi\sigma k h, \qquad (6\text{-}4)$$

which can again be written as an integral

$$B(\sigma) = \int_0^\infty I(\delta) R_h(\delta) D(\delta) \sin 2\pi\sigma\delta \, d\delta, \qquad (6\text{-}5)$$

where $D(\delta)$ is a rectangular function, of width $2\delta_{max}$.

The expression $B(\sigma)$ is the convolution of the sine FT of the odd function $I(\delta)\times R_A(\delta)$ by the cosine FT of the even function $D(\delta)$, which we shall designate by $f(\sigma)$*.

$$B(\sigma) = B_o'(\sigma) * f(\sigma). \qquad (6\text{-}6)$$

The $B(\sigma)$ is again a periodic spectrum with $1/h$ periodicity which is the convolution of the previously considered periodic spectrum by $f(\sigma)$ (Figures 6-4 and 6-5).

The function $f(\sigma)$ is the instrumental line shape (ILS). It is the response of the complete system (interferometer+computer) to a monochromatic line. Just as for classical spectrometers one can define an instrumental resolution $\delta\sigma$.

In the case defined by (6-5)

$$f(\sigma) = \frac{\sin 2\pi\sigma\,\delta_{max}}{2\pi\sigma\,\delta_{max}},$$

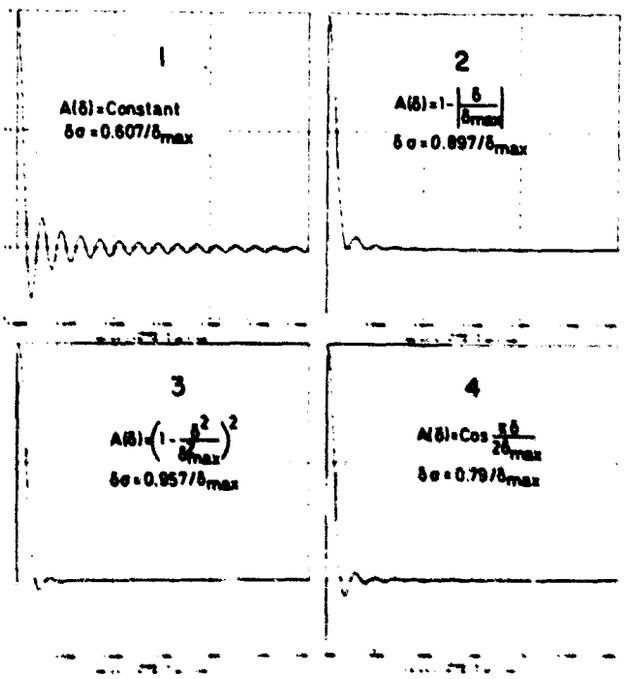


Figure 6-6. Instrumental Line Shape for Several Apodizations. ($A(\delta)$ is the weighting function; $\delta\sigma$ is the half intensity width)

* The result is the same if $I(\delta)$ is even, in this case one performs the cosine FT of the product of two even functions $I(\delta)R_A(\delta)$ and $D(\delta)$.

and we shall use the simple definition

$$\delta\sigma = \frac{1}{2\delta_{max}},$$

which does not differ too much from the usual half maximum intensity criterion.

This first ILS has very strong secondary maxima which can easily be attenuated. One has to weight the interferogram by a suitable "apodization" function $A(\delta)$. Figure 6-6 gives examples of some ILS. In the general case one tries to reduce secondary maxima at the cost of some widening of the central peak[1] (Figures 6-7 and 6-8).

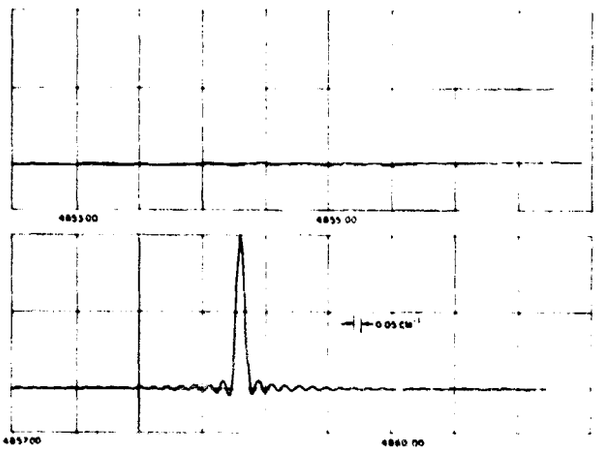


Figure 6-7. Actually Recorded ILS, Without Apodization

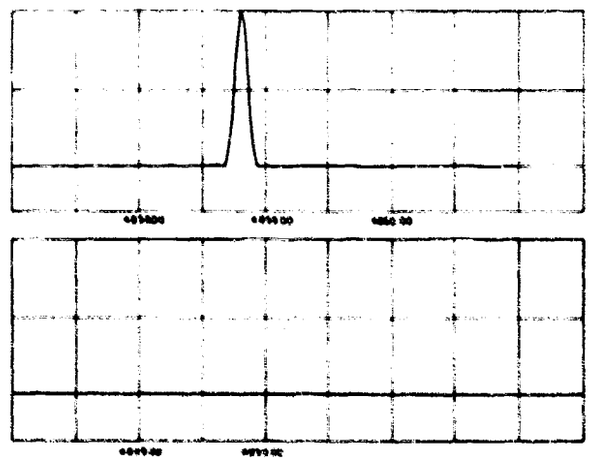


Figure 6-8. Same ILS, with Apodization

$$A(\delta) = (1-\delta^2/\delta_{max}^2)^2,$$

Recorded With He Line 2.05μ, Measured Resolution 0.119 cm^{-1}

In the language we have been using for interpreting (6-5) we say that the ILS scans a periodic spectrum. However, one can also associate the terms within the integral in a different way and write

$$B(\sigma)=B_0(\sigma)*F(\sigma). \tag{6-7}$$

The term $F(\sigma)$ is the cosine FT of the product $R_h(\delta)\times D(\delta)$

$$F(\sigma)=f(\sigma)*R_{1/h}(\sigma). \tag{6-8}$$

This means the ILS is periodic, with period $1/h$, Figure 6-9.

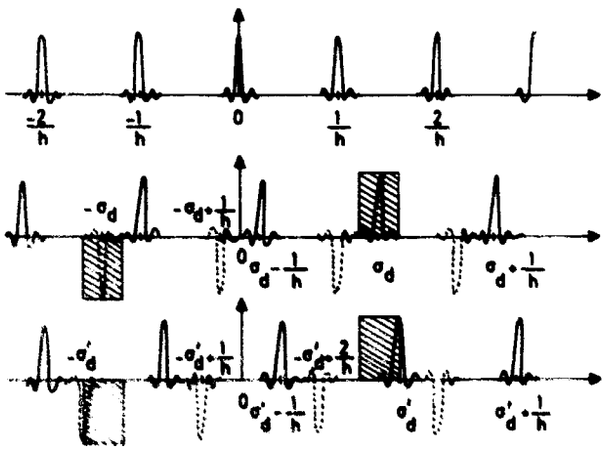


Figure 6-9. Periodic ILS

If the spectrum contains only one line of wavenumber σ_0 the reconstructed spectrum can be written (Figure 6-9)

$$B(\sigma)=\tfrac{1}{2}[f(\sigma-\sigma_0)-f(\sigma+\sigma_0)]. \tag{6-9}$$

One finds two peaks centered on wavenumbers σ_0 and $-\sigma_0$ respectively.

In the case of an extended spectrum, and for $\sigma=\sigma_d$, B can be written

$$B(\sigma_d)$$

$$=\tfrac{1}{2}\int_{-\infty}^{+\infty} B_e(\sigma)[F(\sigma_d-\sigma)-F(-\sigma_d-\sigma)]\,d\sigma. \tag{6-10}$$

The two portions of the spectrum are scanned (Figure 6-9) by two ILS centered on σ_d and $-\sigma_d$, respectively. In order to reconstruct the spectrum without any ambiguity only one peak of the ILS must fall in the two domains. Thus, we are led exactly to the same conditions for selecting h.

If we have infinite resolution, the computed spectrum is identical to the original spectrum; it is strictly zero outside range $\Delta\sigma$, and $B(\sigma_d)$ gives the spectral intensity for σ_d without any approximation. Strictly speaking, since resolution is finite in any actual case, $B(\sigma)$ is unlimited and the first member of (6-10) is in fact

$$B(\sigma_d\pm q/h).$$

In practice, if suitable apodization is used, the spectrum becomes rapidly negligible outside $\Delta\sigma$; thus we shall keep $B(\sigma)$ as it is written in (6-10). Thus we have two periodic functions $B(\sigma)$ and $I(\delta)$ (limited to δ_{max}) which are rigorously Fourier transforms of each other and we can, again without approximation, compute one from the FT of the other, using only a finite number of samples. For $B(\sigma)$ the sampling interval is $\delta\sigma=\dfrac{1}{2\delta_{max}}$ and for $I(\delta)$ it is $h=1/2\,\Delta\sigma$.

Within the free spectral range $\dfrac{1}{2h}$ one will compute $N=\Delta\sigma/\delta\sigma$ independent spectral elements. Thus, N is the minimum number of points needed to describe the interferogram $I(\delta)$ in the noiseless case. Let us consider an actual interferogram, with noise, recorded during total time T, and the resultant spectrum. We showed 10 years ago that the optimum spectrum S/N ratio will be realized if the interferogram has been recorded by the stepping method; ideally, the integration time for each sample is then T/N. With any other method the optimum S/N ratio can only be reached by oversampling, which is permissible only if N is not too large.[6] In the rest of this paper N will always designate the *minimum* number of samples.

6-3 COMPUTATION OF THE SPECTRUM FROM THE INTERFEROGRAM $I(\delta)$. APPLICATION OF THE COOLEY-TUKEY METHOD. INTERPOLATION WITHIN THE SPECTRUM

The classical method for computing the sum (6-3) involves the product of the square matrix of order N by a column matrix with N terms; the computation time is proportional to N^2. The coefficient is a function of the method for getting the cosines and, of course, of the computer itself. The shortest times realized were given by

$$T = 0.8\ 10^{-3} N^2 \text{ for the IBM 704}$$

and

$$T = 0.3\ 10^{-3} N^2 \text{ for the IBM 7040,}$$

where T is expressed in seconds. The highest N actually transformed in this manner was $N = 12\,000$ (with $T = 12h$) in January 1966 at Meudon Observatory (on a 7040).

At about the same time the technique for recording interferograms had reached a stage where 60 000 samples could profitably be used for immediate astronomical problems. In the longer range we could visualize problems with about 10^6 samples. There seemed to be, however, little hope of computing the transform.

We shall presently show that the fast Fourier transform (FFT) of Cooley and Tukey,[12] plus correct use of 3rd generation computers with direct access auxiliary stores have solved the problem; Fourier spectroscopy should no longer be considered an insatiable consumer of digital computer time, a fact which is important for the future of the technique. The computations must be performable within the normal operating conditions of any computing center.

One should first notice that the FFT algorithm in its usual form can transform an arbitrary complex function sampled at $2N = 2^n$ points. If applied directly to our Fourier spectroscopy case, the input complex function will have a *real part* made up of $2N$ interferogram samples. The ones that should have been recorded by going from $-\delta_{max}$ to $+\delta_{max}$, and an *imaginary part* having $2N$ values all equal to zero.

But the interferogram is a real function; furthermore it is either even or odd. There is no need to treat it as an arbitrary function.[10,11] By adapting the Cooley-Tukey algorithm to the case of a real odd function we have reduced the storage space by a factor of 4 and the computation time by 2.8.

6-3.1 Principle of the "Decimation in Time" FFT Method; FFT of an Arbitrary Complex Function

The first description of the technique by Cooley and Tukey[12] goes back to 1965; much literature has been published on the subject since [13-19] and also on its application to Fourier spectroscopy.[20,21]

A new satisfactory approach has been given by Brigham *et al.*[22] We shall here explain the so called "decimation in time" technique[23] because it is very clear and leads to easy explanations of several programs for computing

(a) the FFT of a real function,

(b) the FFT of a real and odd or even function, and

(c) the FFT when N is larger than the central memory size.

Let us call I_k the arbitrary complex function to be transformed with $k \in (0, 1, \ldots, 2N-1)$. In our case the samples I_k would be those of the interferogram recorded from $-\delta_{max}$ to $+\delta_{max}$. We shall call B_r the DFT of the I samples, with $r \in (0, 1, \ldots, 2N-1)$; the samples of B are spread within the range $1/h = 2\Delta\sigma$. In Fourier spectroscopy B is the spectrum.

With the usual notations B_r is the discrete Fourier transform (DFT) of I

$$B_r = \sum_{k=0}^{k=2N-1} I_k e^{\frac{2\pi i r k}{2N}}$$

If we have

$$W = e^{\frac{2\pi i}{2N}} = e^{\frac{\pi i}{N}}, \tag{6-11}$$

then

$$B_r = \sum_{k=0}^{k=2N-1} I_k W^{rk}. \tag{6-12}$$

The entire computation of B from I rests on the following rule

Rule No. 1

The DFT of any given complex function is a linear combination of the two DFT of two functions with half as many terms and issued from the first.

Let Y_k and Z_k be the sets corresponding to the *even* and *odd* values of index k

$$I_k \quad \begin{cases} Y_k = I_{2k} \\ Z_k = I_{2k+1} \end{cases}$$
$$k \in (0, 1, \ldots, 2N-1) \quad k \in (0, 1, \ldots, N-1).$$

Let C and D be the DFT of Y_k and Z_k, respectively, each having only N terms.

Applying (6-11) to them results in

$$C_r = DFT(Y_k) = \sum_{k=0}^{k=N-1} Y_k e^{\frac{2\pi irk}{N}} = \sum_{k=0}^{k=N-1} Y_k W^{2rk},$$

$$D_r = DFT(Z_k) = \sum_{k=0}^{k=N-1} Z_k W^{2rk},$$

$$r \in (0, 1, \ldots, N-1).$$

In (6-12) we separate the *odd* and *even* terms of I and get

$$B_r = \sum_{k=0}^{k=N-1} Y_k W^{2rk} + \sum_{k=0}^{k=N-1} Z_k W^{(2k+1)r} \qquad (6\text{-}13)$$

Then $B_r = C_r + W^r D_r$.

Functions C_r and D_r are periodic of equal period N. Hence

$$C_{N+r} = C_r, \; D_{N+r} = D_r \text{ and } W^{r+N} = -W^r.$$

We can obtain the two sets of samples B from 0 to $N-1$ and from N to $2N-1$ with the two linear combinations

$$\begin{cases} B_r = C_r + W^r D^r \\ B_{N+r} = C_r - W^r D^r. \end{cases} \quad r \in (0, 1, \ldots, N-1) \qquad (6\text{-}14)$$

It follows that we can compute the $2N$ complex samples of the DFT of the entire interferogram I from the N complex samples of the DFT of each of the two subsets Y and Z. If we apply the same reasoning $\log_2 2N$ times we can obtain B, DFT of I with only $2N \log_2 2N$ elementary operations.

Figures 6-10 and 6-13 present the case where $N_d = 8 = 2^3$. Let us use the symbol P for the DFT operation. Figure 6-10 illustrates the classical DFT, giving the 8 B outputs from the 8 I inputs. In Figure 6-10 the 8 values of I have been sorted into 2 subsets P_1 and P_2, corresponding to the even and odd indices of I. To each of these subsets a DFT is applied and the operations described by (6-14) give the final 8 B outputs from the two partial DFT.

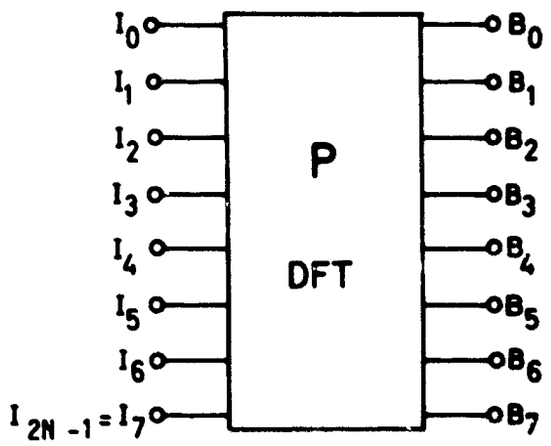


Figure 6-10. Discrete Fourier Transform of 8 Input Points

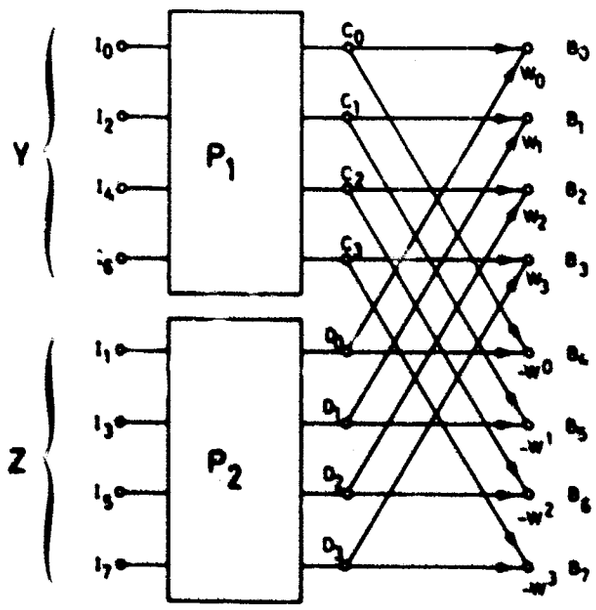


Figure 6-11. Computation of the Sampled Spectrum After Separating Set P into Subsets P_1 and P_2

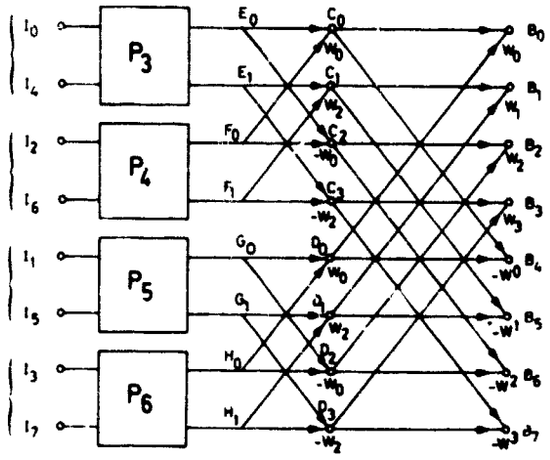


Figure 6-12. Separation of Sets P_1 and P_2 into P_3, P_4 and P_5, P_6

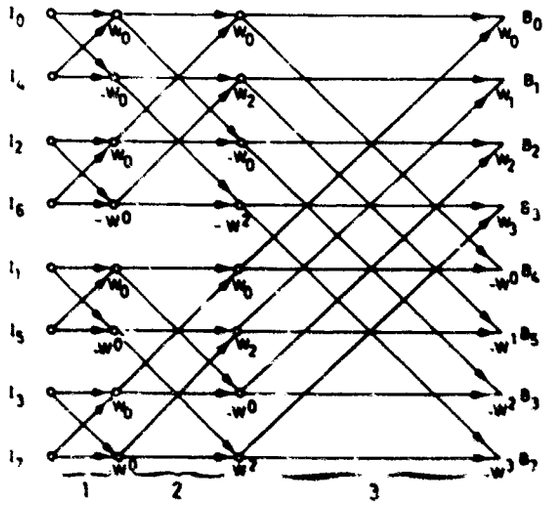


Figure 6-13. Complete Diagram of "Decimation in Time" FFT Technique

A new subdividing of P_1 in P_3, P_4 and P_2 in P_5, P_6 again gives the final 8 B_i but operations (6-14) have to be applied twice (Figure 6-12). One more step, in which P_3, P_4, P_5, P_6 are divided, and we get Figure 6-13 in which are presented all the elementary operations of the FFT in the case $N = 8$.

In the general case we must at each stage:

(1) divide in 2 parts the subsets given by the previous operation, hence we have $\log_2 2N$ stages; and

(2) combine the 2 results of the DFT operation applied to the new subsets in order to recover the DFT results of the preceding stage. This combination implies 1 multiplication and 1 addition. Hence the total number of operations is proportional to $2N \log_2 2N$.

The time reduction compared with the classical technique is of the order of $2N/\log_2 2N$ *provided* one has free use of $4N$ locations in the main random access storage. For instance:

if $N = 2^{10} = 1024$, the gain is 100;

and if $N = 2^{20} \approx 10^6$, the gain is 50 000.

6-3.2 Programming the Algorithm

Looking at Figure 6-12 we can foresee the difficulties we will meet when programming the algorithm.

6-3.2.1 INITIAL SORTING OF THE DATA

Before the first stage of computation the data must be sorted and put into a different order. There are two possible solutions:

6-3.2.1.1 Inverting the Bits

In the $N = 4$ example let us write in binary form (with $\log_2 2N$ bits) the order number of the interferogram samples, and invert the bits

	BI
$I_0 = 000$	$000 = 0$
$I_1 = 001$	$100 = 4$
$I_2 = 010$	$010 = 2$
$I_3 = 011$	$110 = 6$
$I_4 = 100$	$001 = 1$
$I_5 = 101$	$101 = 5$
$I_6 = 110$	$011 = 3$
$I_7 = 111$	$111 = 7$

We find precisely the order in which the data samples must be put before performing the first operation.

6-3.2.1.2 Recurrence Formula

One can also use a recurrence formula which will give the indices i. For instance, in order to get $i(n)$ we must determine P according to

$$2^P < n \leq 2^{P+1}$$

Then

$$i(n) = i(n-2^P) + \frac{2N}{2^{P+1}}, \qquad (6\text{-}15)$$

with $i(1)=0$,

$$i(2)=2N/2=N,$$

$$i(3)=2N/4=N/2.$$

6-3.2.2 LIMITATIONS DUE TO SIZE OF CENTRAL STORAGE

If we look at the last stages of the computation, we see that the two numbers involved in one given combination do not follow each other. Thus, if we are to apply the algorithm directly, we must keep the entire computation within the central random access memory.

Let M be the memory size. In any computer a fraction of it will always be used up by the system and the program. Since M is in most cases a power of 2, only $M/2$ will remain available for storing the $2N$ complex numbers, which means $4N \leq M/2$ and $N \leq M/8$. For instance with $M = 32\ K$ words* we can test at the most an interferogram with $N = 4096$ samples.

The conclusion is thus the same one we reached about the interferometer. The greatest difficulties will be met with large numbers of samples, when $N > N_L$, if we put $N_L = M/8$. Then we must use nonrandom access auxiliary storage. This case will be studied later.

6-3.3 Real Functions

In the case of a real function I with $2N$ samples we would like to treat only $2N$ samples, that is, compute the FFT of complex function E with N samples made up by suitable intercombinations of the $2N$ real samples I. Then B will be reconstructed using a linear association of the $E(\delta)$ FFT results. Both rule 1 and a new rule 2 are needed.

Rule No. 2:

Only $2N$ memory words are needed to store within the central memory the N complex samples of a complex function and the N complex samples of the conjugate function.

6-3.3.1 PRINCIPLE

Let E_k be a complex function, with

$$E_k = Y_k + iZ_k, \qquad (6\text{-}16)$$

and F_r the complex DFT of E_k computed from N complex samples of E_k. We note that C_r and D_r have already been written as

$$C_r = \text{DFT}\ (Y_k), \text{ and } D_r = \text{DFT}\ (Z_k).$$

$$F_r = C_r + iD_r. \quad r \in (0, 1, 2, \ldots, N-1) \qquad (6\text{-}17)$$

The Y_k and Z_k are real; C_r and D_r are complex and Hermitian (see Appendix A), which means:

$$C_r = C^*_{-r}, \text{ and } D_r = D^*_{-r}, \qquad (6\text{-}18)$$

and we obtain

$$F^*_{-r} = C^*_{-r} - iD^*_{-r} = C_r - iD_r. \qquad (6\text{-}19)$$

Comparison between (6-17) and (6-19) gives:

$$C_r = \frac{F_r + F^*_{-r}}{2},$$

$$D_r = \frac{F_r + F^*_{-r}}{2i}. \qquad (6\text{-}20)$$

We shall obtain the $2N$ complex samples of B by the usual association (6-14) thus:

$$B_r = \frac{F_r + F^*_{-r}}{2} + W^r \frac{F_r - F^*_{-r}}{2i}$$

$$B_{N+r} = \frac{F_r + F^*_{-r}}{2} - W^r \frac{F_r - F^*_{-r}}{2i} \qquad (6\text{-}21)$$

6-3.3.2 NOTES FOR PROGRAMING

When F is periodic and its period is equal to N, we have

$$F^*_{-r} = F^*_{N-r};$$

* $1K = 2^{10} = 1024$.

I_k is a real function sampled at $2N$ points. Thus, B is complex and Hermitian. Consequently,

(a) $B_{N+r}=B^*_{N-r}$; (6-22)

(b) the real part of B_r for $r \epsilon(0, 1, 2, \ldots, 2N-1)$ is an even function of $2N$ samples making 2 sets symmetrical with respect to $r=N$. The imaginary part is an odd function of $2N$ samples with respect to the same $r=N$.

Thus, in order to have the entire spectrum it is sufficient to store N complex values of B, which can be either:

$$B_r=\tfrac{1}{2}[(F_r+F^*_{N-r})-iW^r(F_r-F^*_{N-r})], \quad (6\text{-}23)$$

or

$$B_{N+r}=B^*_{N-r}$$
$$=\tfrac{1}{2}[(F_r+F^*_{N-r})+iW^r(F_r-F^*_{N-r})]. \quad (6\text{-}24)$$

From one of the two sets it is easy to calculate the other. In the case of a perfectly even or odd interferogram, the spectrum contains only N real numbers which are either the real or the imaginary part of B_r. One must select either the *first* half of the $2N$ samples of the real part or of the imaginary part of B_r, or the *second* half according to whether the optical spectrum is located on the first or second half of range $1/h$.

6-3.4 Odd Real Functions

6-3.4.1 PRINCIPLE

The interferogram is not merely real but also either odd or even. Thus, we are led to use this additional property and use a technique by which the FFT will be applied to N samples only,[11] instead of $2N$. From these N samples, which characterize one half of the odd or even function, we can construct an Hermitian function H_k having N complex samples. The DFT of H_k is a real function which can be obtained by linear combinations of the results of a DFT performed on an intermediate function F_k with $N/2$ complex terms only.

We shall treat the odd function case here (the even function case is treated in Appendix B).

$$I_k=-I_{2N-k}, \quad k \epsilon(0, 1, \ldots, 2N-1) \quad (6\text{-}25)$$

$$I_{(N+k)}=-I_{(N-k)}.$$

We have previously defined E_k by (6-16) from Y_k and Z_k. In the same way:

$$G_k=I_{2k+1}+iI_{2k}, \quad k \epsilon(0, 1, \ldots, N-1)$$

and we define the new function H_k by

$$H_k=\tfrac{1}{2}(G_k+G^*_{N-k}),$$
$$G^*_{N-k}=+I_{2(N-k)+1}-iI_{2(N-k)}$$
$$=I_{2N-2k+1}-iI_{2N-2k}. \quad (6\text{-}26)$$

Or, making use of (6-25)

$$G^*_{N-k}=-I_{2k-1}+iI_{2k},$$

and

$$H_k=\tfrac{1}{2}[(I_{2k+1}-I_{2k-1})]+iI_{2k}. \quad (6\text{-}27)$$

Function H_k is Hermitian.* It is made up of N complex terms. Its DFT $K(r)$ is a real function of N terms from which we can obtain the N real spectrum samples by a linear combination.

We shall see in the next paragraph how this DFT can be reconstructed from the DFT of an auxiliary

* Let us compute H_{k1} and H_{k2} for 2 points symmetrical with respect to $H_{N/2}$

$$k_1=\frac{N}{2}-p \quad \text{and} \quad k_2=\frac{N}{2}+p, \; p \epsilon\left(0, 1, \ldots, \frac{N}{2}-1\right)$$

$$H_{k1}=\tfrac{1}{2}[I_{N-2p+1}-I_{N-2p-1}]+iI_{N-2p},$$
$$H_{k2}=\tfrac{1}{2}[I_{N+2p+1}-I_{N+2p-1}]+iI_{N+2p}.$$

Making use of (6-20)

$$I_{N+2p}=-I_{N-2p},$$
$$I_{N+2p+1}=-I_{N-2p-1},$$
$$I_{N+2p-1}=-I_{N-2p+1},$$
$$H_{k2}=\tfrac{1}{2}[I_{N-2p+1}-I_{N-2p-1}]-iI_{N-2p}.$$

The real part of H_k is even and the imaginary part is odd; thus, it is Hermitian.

$$H^*_{k2}=H_{k1}.$$

function having only $N/2$ complex terms.

$$K_r = \sum_{k=0}^{k=N-1} H_k W^{2kr},$$

or, replacing H_k,

$$K_r = \tfrac{1}{2} \sum_{k=0}^{k=N-1} (I_{2k+1} - I_{2k-1}) W^{2kr} + i \sum_{k=0}^{k=N-1} I_{2k} W^{2rk}. \quad (6\text{-}28)$$

Let us consider the term

$$K'_r = \sum_{k=0}^{k=N-1} I_{2k-1} W^{2kr},$$

and put $l=k-1$. Then

$$K'_r = \sum_{l=-1}^{l=N-2} I_{2l+1} W^{2rl} W^{2r}.$$

Since the I_{2k+1} function is periodic with periodicity N the sum over a complete period will not be changed by shifting the limits. Thus:

$$K'_r = \sum_{l=0}^{l=N-1} I_{2l+1} W^{2rk} W^{2r}.$$

Then (6-28) becomes

$$K_r = \tfrac{1}{2}(1 - W^{2r}) \sum_{k=0}^{k=N-1} I_{2k+1} W^{2rk} + i \sum_{k=0}^{k=N-1} I_{2k} W^{2rk}, \quad (6\text{-}29)$$

from which we get

$$K_{N-r} = \tfrac{1}{2}(1 - W^{-2r}) \sum_{k=0}^{k=N-1} I_{2k+1} W^{2rk} + i \sum_{k=0}^{k=N-1} I_{2k} W^{-2rk}. \quad (6\text{-}30)$$

Since K is real, $K_{N-r} = K^*_{N-r}$, and (6-30) becomes equivalent to (6-31).

$$K_{N-r} = \tfrac{1}{2}(1 - W^{2r}) \sum_{k=0}^{k=N-1} I_{2k+1} W^{2rk} - i \sum_{k=0}^{k=N-1} I_{2k} W^{2rk}, \quad (6\text{-}31)$$

in which the exponentials are the same as in (6-29).

On the other hand the DFT B_r of I_k can be written

$$B_r = \sum_{k=0}^{k=2N-1} I_k W^{rk}, \quad (6\text{-}32)$$

or

$$B_r = \sum_{k=0}^{k=N-1} I_{2k+1} W^{2rk} W^r + \sum_{k=0}^{k=N-1} I_{2k} W^{2rk}. \quad (6\text{-}33)$$

From (6-29, 6-32, 6-33) we can compute the N real samples of

$$B_r = W^r \frac{(K_r + K_{N-r})}{1 - W^{2r}} + \frac{1}{2i}(K_r - K_{N-r})$$

$$= \frac{K_r + K_{N-r}}{W^{-r} - W^r} - \frac{i}{2}(K_r - K_{N-r}),$$

from which we get finally

$$B_r = \frac{i}{2}\left[\frac{(K_r + K_{N-r})}{\sin\frac{\pi r}{N}} - (K_r - K_{N-r})\right]. \quad (6\text{-}34)$$

This is a pure imaginary function, the N samples of which are obtained by linear combinations of N real samples of K_r. The transform K_r is the FT of the Hermitian function H_k of which we have stored $N/2$ complex samples.

6-3.4.2 FT OF THE HERMITIAN FUNCTION H_k

We have seen above that the FT of a real function is Hermitian. Inversely, the FT of an Hermitian function is real. By following the same path in the opposite direction we can define a new function F_k with $N/2$ complex terms, related to $N/2$ complex terms of H_k by the relationship

$$F_k = \tfrac{1}{2}\left[(H_k + H^*_{\frac{N}{2}-k}) + iW^{-2rk}(H_k - H^*_{\frac{N}{2}-k})\right]. \quad (6\text{-}35)$$

The DFT of the $N/2$ terms of function F_k is a complex function E_r with $N/2$ terms also, as

$$E_r = Y_r + iZ_r. \quad r \in \left(0, 1, \ldots, \frac{N}{2}-1\right)$$

The terms Y_r and Z_r are, respectively, the even or odd samples of the real function K_r:

$$Y_r = K_{2r}, \text{ and } Z_r = K_{2r+1}.$$

Thus, the N terms of function K_r are found.

6-3.4.3 SUMMARY

The three procedures for treating (a) complex, (b) real, and (c) real and odd (or even) data can be symbolized in the following way:

(a) Complex
Real part:
I_k (0, 1, . . . , $N-1$) followed by N zeros imaginary part $2N$ zeros → [FFT, $2N$ complex] → B_r $2N$ complex ;

(b) Real
I_k (0, 1, . . . , $N-1$) followed by N zeros $2N$ samples → E_k N complex samples → [FFT, N complex] → F_r N complex → B_r $2N$ complex ; (Hermitian)

(c) Real and Odd (or even)
I_k (0, 1, . . . , $N-1$) N samples → H_k Hermitian N complex samples → F_k $\frac{N}{2}$ complex → [FFT, $\frac{N}{2}$ complex] → E_r → K_r N real samples → B_r real N samples .

In the real function case the FFT involves only N complex samples instead of $2N$ complex samples; in the real and odd (or even) case only $N/2$. The entire process, however, involves additional sine operations. Thus, while the gains in storage space are actually 2 and 4, the time reductions so far are only 1.5 and 2.8.

6-3.5 Fourier Transform with a Large Number of Samples

6-3.5.1 PRINCIPLE

If the number of interferogram samples N is greater than N_L (with $N_L = M/4$), the computation can no longer be performed within the central storage alone and one must use auxiliary direct access memories, that is, drums or disks (as opposed to sequential access magnetic tapes).[26]

We shall now give the principle, by using an extension of the rules, which are apparent from Figure 6-11.

Just as we had called N_L the maximum number of interferogram samples that could be treated within the central storage, let us call $2N_L$ the maximum number of complex samples in the complex function we now want to transform. We must sort the data into the minimum possible number of subsets P; each must be transformable by FFT within the central

storage. Thus, each must contain $2N=2N_L$ samples and we shall have to perform $K=N/N_L$ FFT operations. The term K will always be a power of 2.

Once K is determined, two problems remain:

(a) How to sort the primary samples into the subsets P? This is solved according to the following rule. Let us give each subset P an index varying from 0 to $K-1$, consider the binary formulation with $\log_2 K$ bits of these indices in their natural order, and then invert the bits. A new sequence of indices is obtained; these are found to be the indices J of the first sample in each subset P; and the following samples will be found to have the indices $J+K$, $J+2K, \ldots, J+(2N-1)K$. For example, let us treat the $K=8$ case. Inverting the binary representation of numbers 0, 1, . . . , 7 is given.

We conclude that the first sample in each subset has index 0, 4, 2, 6, 1, 5, 3, 7.

(b) How many recombinations will be needed after the FFT operations? Since K is a power of 2 we can write $K=2^p$, and we shall have p stages of recombinations from the partial FT operations in order to get the $2N$ final samples in their correct sequence. In the just mentioned example we should have 3 stages.

6–3.6 Results

The computation times actually obtained are summarized in Table 6-1 and Figure 6-14 both for classical FT and FFT as performed by six different computers (IBM 704, IBM 7040, IBM 360/75, UNIVAC 1108, IBM 1130, CDC 6600). They show the several orders of magnitude of reductions in computing time realized in the last ten years both by better programing techniques and by the availability of faster computers. Skillful use of direct access auxiliary storage in third generation computers has been particularly important.

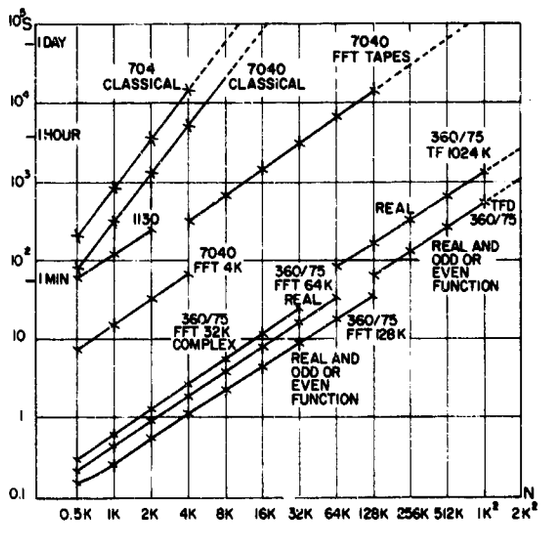


Figure 6-14. Diagram of Computation Times. Crosses indicate actually measured times; dotted curves are extrapolations

The computation times are given by the following formulas:

(1) IBM 7040 ($M=32\,K$) Program FFT 4 K (for $N\leq 4\,K$, complex function).*

$$T=0.428\times10^{-3}\,2N\log_2 2N+0.166\times10^{-2}\,2N.$$

(2) IBM 360/75 ($M=256\,K$) Program FFT 32 K (for $N\leq 32\,K$).†

$$T=0.152\times10^{-4}\,2N\log_2 2N+0.145\times10^{-3}\,2N.$$

(3) UNIVAC 1108 ($M=64\,K$) Program FFT 8 K (for $N\leq 8\,K$, complex function).**

$$T=0.373\times10^{-4}\,2N\log_2 2N+0.261\times10^{-3}\,2N.$$

(4) IBM 360/75 ($M=256\,K$) Program FFT 64 K ($N\leq 64\,K$, real function).††

$$T=0.158\times10^{-4}N\log_2 N+0.206\times10^{-3}\,2N.$$

CDC 6600 ($M=128\,K$) (real function)

$$T=0.17\times10^{-4}N\log_2 N+0.230\times10^{-3}\,2N.$$

IBM 360/75 ($M=256\,K$) Program FFT 128 K ($N\leq 128\,K$, real and odd or even function).

For $N=8\,K$, T is almost proportional to N,
$T=0.15\times10^{-3}\,N$.

IBM 1130 ($M=32\,K$) Program in double precision

T is practically proportional to N,
$T=0.11\,N$.

* In the first usable FFT programs sorting of the numbers was accomplished by bit inversion which introduced in the computing time a term proportional to N^2. This was negligible for small N only. Later on we eliminated this term which was replaced by one proportional to N; this implied using N additional storage locations.

† The time gain realized by simply using the 360/75 instead of the 7040 was found equal to 23.

** For this particular problem the 360/75 is 1.8 times faster than the 1108.

†† In this case the formula involves N instead of $2N$.

Table 6-1. Computation Times for Different Computers and Programs; N is the Number of Interferogram Samples from 0 to δ_{max}

COMPUTER	704	7040 32 K	7040 32 K	7040 32 K	360/75 256 K	UNIVAC 1108 64 K	360/75 256 K	1130 32 K	6600 CDC 128 K	360/75 256 K	360/75 256 K	360/75 256 K
PROGRAMS	classical	classical	FFT 4 K	with magnetic tapes	FFT 32 K $N \leq 32$ K	FFT 8 K $N \leq 8$ K	FFT 64 K $N \leq 64$ K	FT 1130 $N \leq$ double precision	FT 6600 $N \leq$	FFT 128 K $N \leq 128$ K	TF 1024 K with disks	TF 2048 K with disks
FUNCTION	real	real	complex	complex	complex	complex	real	real	real	real and even or odd	real	real and even or odd
MAIN STORAGE AREA	$4N$	$4N$	$5N$	$5N$	$5N$	$5N$	$5N/2$	$5N/2$	$5N/2$	$3N/2$	$5N/32$	$3N/32$
$N=0.5$ K	3 min 30 sec	78 sec	7.22 sec		0.30 sec	0.54 sec	0.22 sec	60 sec	0.212 sec	0.12 sec		
$N=1$ K	14 min	5 min 14 sec	15.33 sec		0.66 sec	1.18 sec	0.44 sec	120 sec	0.450 sec	0.26 sec	3.7 sec	
$N=2$ K	55 min 55 sec	20 min 58 sec	32.45 sec		1.34 sec	2.40 sec	0.90 sec	240 sec	0.950 sec	0.50 sec	5.18 sec	
$N=4$ K	3h 43 min	1h 23 min	68.41 sec	5 min 15 sec	2.88 sec	5.20 sec	1.85 sec		2.00 sec	1.10 sec	8.14 sec	
$N=8$ K	14h 56 min	5h 35 min		9 min 30 sec	5.85 sec	10.5 sec	3.82 sec		4.204 sec	2.46 sec	13.32 sec	
$N=16$ K				21 min	12.22 sec		7.86 sec		8.64 sec	5.09 sec	22.94 sec	
$N=32$ K				50 min	25.45 sec		16.00 sec		17.6 sec	9.93 sec	42.92 sec	
$N=64$ K				1h 45 min			33.44 sec			19.97 sec	1 min 24 sec	
$N=128$ K				3h 32 min			39.42 sec			39.42 sec	2 min 46 sec	1 min 8 sec
$N=256$ K											5 min 27 sec	2 min 16 sec
$N=512$ K											11 min	4 min 32 sec
$N=1024$ K											22 min 8 sec	9 min 5 sec

6-3.6.1 IBM 7040 ($M=32\,K$) PLUS MAGNETIC TAPES

This program was evolved specially for the $N>4\ K$ complex function case on the 7040. The algorithm had to be modified in order to make use of 4 magnetic tapes on which sequentially ordered files were stored.

$$T=0.228\times10^{-2}\,N_d\log_2 2N+0.747\times10^{-2}\,2N.$$

Thus, computation time was 5 times longer than in the purely central storage case with the same computer. This program was used for computing all the planetary spectra later published in atlas form[25] for most of which N was of the order of 60 000. This program could have been much improved, but this was made unnecessary by the availability of the 360/75 with disk stores.

6-3.6.2 IBM 360/75 ($M=256\ K$, PLUS DISKS) PROGRAM FFT 1024 K. (REAL FUNCTION WITH $N\leq1024\ K$)

The first step involves sorting the data in 16 sets, each containing $2N/16=128\ K$ samples, which are stored on disks. During the second step one computes the FT of each of these sets using FFT 64 K. In the third step the 16 sets undergo 4 stages of intercombination, and the global FT is found; the spectral samples (N real and N imaginary) are found in the correct sequence on two different disk files.

The general time is the sum of T_1, T_2, T_3:

$$T_1=0.562\times10^{-4}\,2N,$$

$$T_2=0.158\times10^{-4}\,N\log_2 N+0.206\times10^{-3}\,N_d,$$

$$T_3=0.336\times10^{-3}\,2N.$$

Thus,

$$T=0.158\times10^{-4}\,N\log_2 N+1.119\times10^{-3}\,N,$$

if $N=1024\ K$, $T=22$ min 8 sec.

About $\frac{3}{4}$ of this time is used up by handling of the files which explains why beyond $N=32\ K$ the time is proportional to N. The size and number of files have been optimized for $N=1024\ K$ and are not well suited to $N=64\ K$, in which case one would of course use FFT 64 K, compared with which the time increase is about 3.

6-3.6.3 IBM 360/75 TFD (REAL AND ODD FUNCTION)

The time T is still almost proportional to N;

$$T=0.55\times10^{-3}\,N.$$

In the case of an even or odd interferogram, half of the samples in each subset are equal to zero. The storage size needed is $5\,N/32$ in the central memory and 2.5×10^6 in the disks.

6-3.6.4 IBM 360/75 TFD ($M=256\ K$ PLUS DISKS)

In order to treat the case of a real odd $2048\ K$ function corresponding to a $1024\ K$ interferogram, one uses the Hermitian function H_k (6-26), which has $1024\ K$ complex samples. In order to compute the FT it is sufficient to use only $512\ K$ samples which are subdivided into 8 subsets with $64\ K$ each; it is possible to compute the FT of each subset within the central memory. Then one has only to perform the previously described combinations. The storage size in the central memory stays equal to $5\,N/32$, since one tries to make the best use of the central memory; however, the storage size on the disks is reduced to 1.5×10^6 words and the computation time is reduced by a factor of 2.

6-3.7 Interpolation in the Spectrum

We have stressed the fact that N is the *minimum* number of samples needed to describe the ideal interferogram which is the FT of the spectrum; it follows that N is also the minimum number of samples needed to describe the spectrum—with just one sample per spectral element. They do contain sufficient information to reconstruct the completely smoothed spectrum, but this cannot be done by a simple linear interpolation. We shall have to reconstruct intermediary points by a special type of interpolation. We shall call primary spectral points those that are obtained as a result of the FT operation and secondary points the interpolated ones. We have found that sufficient accuracy for most practical purposes can be realized, with 4 secondary samples between 2 primary ones; then linear interpolation (a standard plotter program) can be applied to draw a smooth curve. In some cases (accurate measurement of line shape or position) more secondary samples are used.[1,5-7,28,29]

6-3.7.1 INTERPOLATION PRINCIPLE

The two periodic functions $B(\sigma)$ and $I'(\delta)$ are Fourier transforms of each other. (Here we take the case of an odd interferogram, but the following discussion applies equally well to the even case.) The period of $I'(\delta)$ around $\delta=0$ is our real interferogram (Figure 6-15*a*).

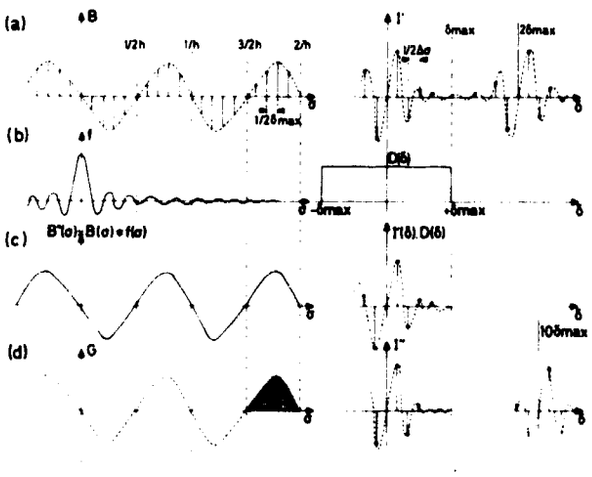


Figure 6-15. Interpolation in the Spectrum (Without Apodization)

The two sampling intervals are $\delta\sigma = 1/2\delta_{max}$ for $B(\sigma)$ and $h = \frac{1}{2\Delta\sigma}$ for $I'(\delta)$. Let us consider the rectangular function $D(\delta)$, which is unity for $-\delta_{max} < \delta < +\delta_{max}$ and 0 outside, and its Fourier transform

$$f(\sigma) = \frac{\sin 2\pi\sigma\,\delta_{max}}{2\pi\sigma\,\delta_{max}},$$

(Figure 6-15*b*), which is the same as the ILS. Let us multiply $I'(\delta)$ by $D(\delta)$. This is equivalent to isolating one of the periods of $I'(\delta)$, thus reconstructing the original interferogram $I(\delta)$. Since the FT of the product of two functions is the convolution of their FT, the FT $B''(\sigma)$ of $I'(\delta) \cdot D(\delta)$ will be (Figure 6-15*c*)

$$B''(\sigma) = B(\sigma) * f(\sigma).$$

This is a continuous function, equal to $B(\sigma)$ for $\sigma = K\,\delta\sigma$, where K is an integer taking all possible values from $-\infty$ to $+\infty$.

This is equivalent to stating that any spectral point of wavenumber σ_0 can be computed in a unique way, without any approximations, by the integral

$$B''(\sigma_0) = \int_{-\infty}^{+\infty} B(\sigma)\,\frac{\sin 2\pi(\sigma_0-\sigma)\,\delta_{max}}{2\pi(\sigma_0-\sigma)\,\delta_{max}}\,d\sigma. \qquad (6\text{-}36)$$

The sampling interval of the interpolated spectrum (that is, the distance between secondary samples) can be arbitrarily small. If we decide to have 4 secondary samples between primary ones, the FT $I''(\delta)$ of the secondary samples spectrum $G(\sigma)$ is a periodic function with period $10\delta_{max}$.

6-3.7.2 APODIZATION OF THE SPECTRUM

Interpolation (6-36) as performed upon primary samples $B(\sigma)$ computed from an unweighted interferogram $I'(\delta)$ gives an unapodized spectrum $B''(\sigma)$. An apodized spectrum can be computed in two different ways:

(1) Instead of starting from $I'(\delta)$ as it is recorded, we multiply it by a weighting function $A(\delta)$; then the primary points $B(\sigma)$ with separation $\delta\sigma = \frac{1}{2\delta_{max}}$ (just as in the previous case) will be new samples of the apodized spectrum. If we again convolve the $B(\sigma)$ samples by the same $f(\sigma)$ function, the continuous apodized spectrum is obtained.

(2) Let us consider again the sampled functions $B(\sigma)$ and $I'(\delta)$ (Figure 6-16*a*) and the weighting function $A(\delta)$ whose FT is $f(\sigma)$ the new wanted ILS (Figure 6-16*b*). The FT of the product $I'(\delta) \cdot A(\delta)$ is (Figure 6-16*c*)

$$B''(\sigma) = B(\sigma) * f(\sigma).$$

This is also a continuous function; however, $B''(\sigma) \neq B(\sigma)$ for $\sigma = K\,\Delta\sigma$; for these same wavenumbers $B''(\sigma)$ is what we should have found by performing the FT of the interferogram weighted for apodization.

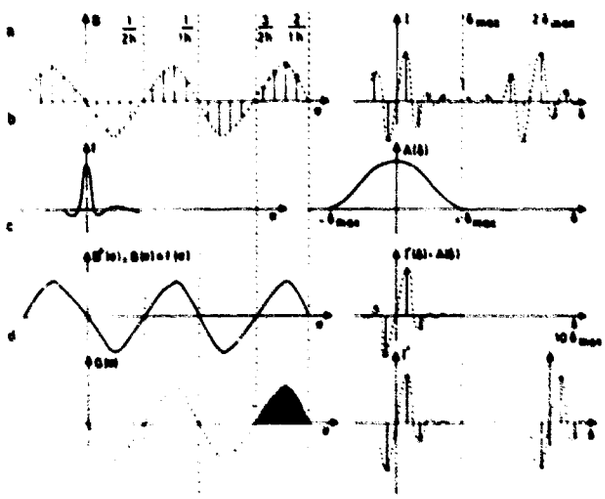


Figure 6-16. Interpolation in the Spectrum (With Apodization)

Any spectral point can be computed by the integral

$$B''(\sigma_0) = \int_{-\infty}^{+\infty} B(\sigma) f(\sigma_0 - \sigma)\, d\sigma. \qquad (6\text{-}37)$$

Again, if the secondary sampling interval is $\delta\sigma/5$, the repetition period of the weighted and sampled interferogram is $10\delta_{max}$ (Figure 6-16*d*).

To summarize: We can either weight the interferogram, compute the FT, and convolve the so obtained apodized primary samples by a pure sinc function, or transform the unweighted interferogram and convolve the unapodized primary samples by the wanted ILS. The second technique is preferable since performing only one FT we can get a given spectrum with several different ILS.

6-3.7.3 NOTES ON PROGRAMMING

In the ideal case described, interpolation involves no approximations and gives the exact result. Computing does, however, introduce an approximation. The integrals (6-36) and (6-37) stretch from $-\infty$ to $+\infty$, and we shall have to limit them to $\pm\sigma_L$.

$$B''(\sigma_0) = \int_{-\sigma_L}^{+\sigma_L} B(\sigma) f(\sigma_0 - \sigma)\, d\sigma. \qquad (6\text{-}38)$$

Cutting off $f(\sigma)$ means the ideal perfect filter $D(\delta)$ is replaced by $D'(\delta)$, which is the result of a convolution between $D(\delta)$ and function $a(\delta)$, FT of $A(\sigma)$, which is the rectangular function cutting off $f(\sigma)$ (Figure 6-17).

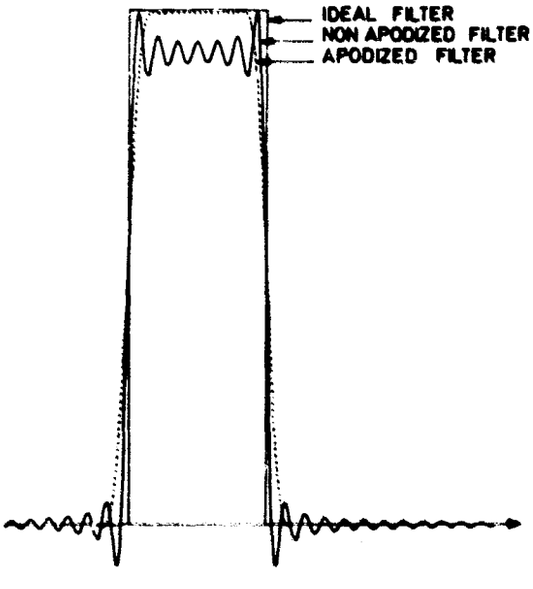


Figure 6-17. Filtering

In order to reduce computing time one has to make σ_L as small as possible, and to improve the filtering of $I'(\delta)$ two essential precautions must be taken.

6-3.7.3.1 Computation of Primary Samples

When computing the transform, if one takes a sampling interval slightly less than the theoretical $\frac{1}{2\delta_{max}}$, for instance $\delta\sigma = 0.98 \times \frac{1}{2\delta_{max}}$, the repetition period of $I'(\delta)$ will be 2.04 δ_{max} (Figure 6-18*a*). Then the transition region of the filter will correspond to a range where $I'(\delta) = 0$ (Figure 6-18*b*). Reducing the sampling interval of $B(\sigma)$ means, in the case of the FFT, filling only 98 percent of the tables intended for the data with actual interferogram samples, and the remaining 2 percent with zeros. For instance, in the case of our longest transforms so far, the 10^6 interferogram samples are followed by zeros up to

$$1024\, K = 1.05 \times 10^6.$$

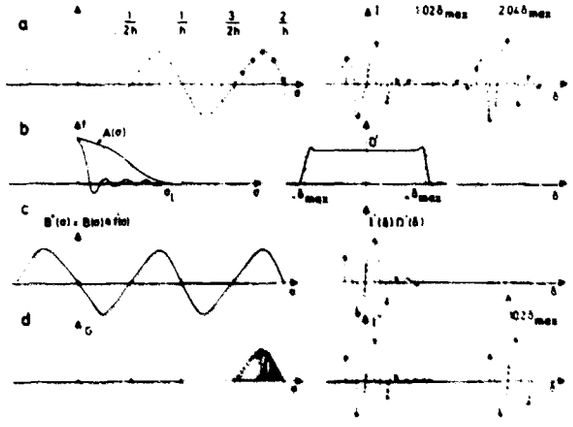


Figure 6-18. Filtering in Spectrum Interpolation

6-3.7.3.2 Weighting of $f(\sigma)$

In actual practice the ILS functions $f(\sigma)$ will not be used directly, but weighted by an "apodizing" function, such as (Figure 6-18*b*)

$$A(\sigma) = (1 - \sigma^2/\sigma_L^2)^2.$$

Then the filter $D''(\delta)$ will be the convolution $D(\delta) * a(\delta)$, where $a(\delta)$ is again the FT of $A(\sigma)$. Oscillations are then considerably damped. There is abundant literature on the subject of filter improvements.

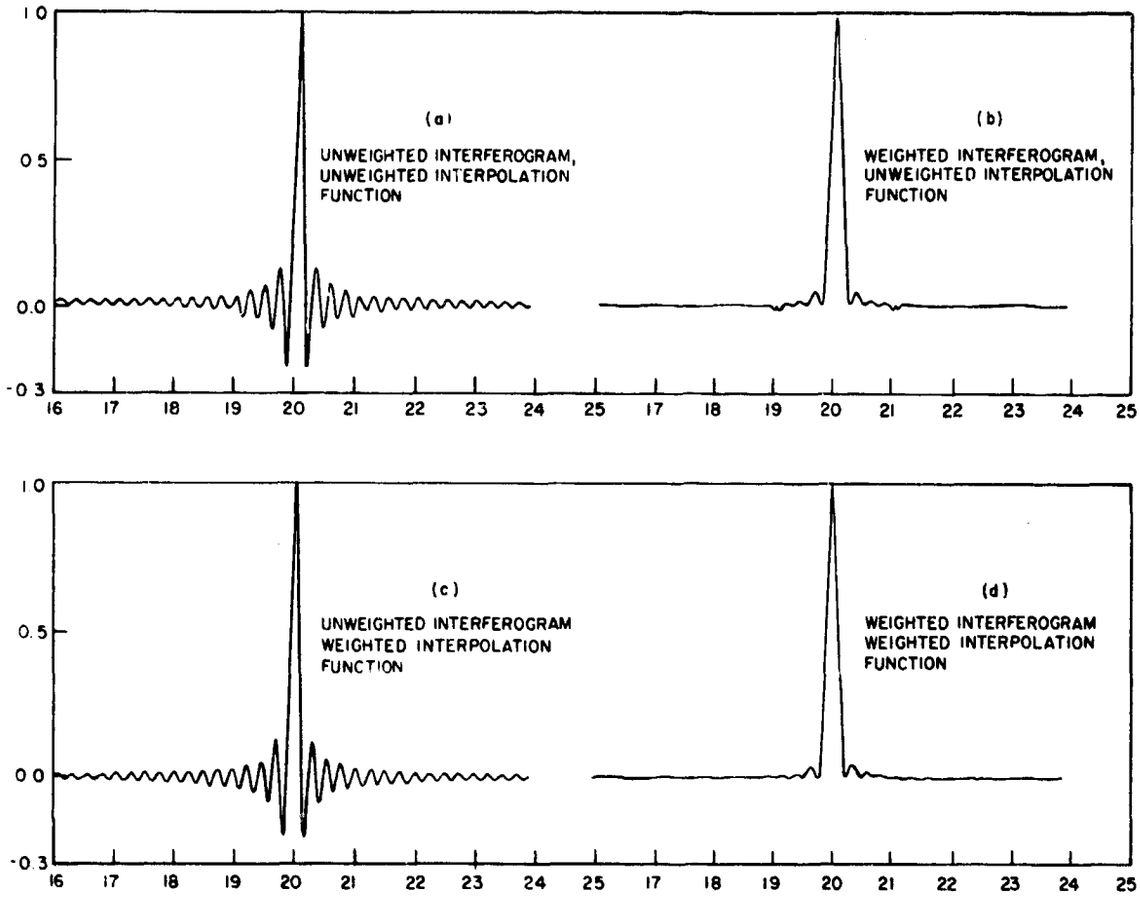


Figure 6-19. Instrumental Line Shape Obtained by Interpolation; Interpolation Function 10 Cycles

Thanks to these two essential precautions we have been able to obtain secondary spectral samples with less than 10^{-5} intensity errors, using only 10 sidelobes of function $f(\sigma)$. If again there are 4 secondary samples between 2 primary ones, this means we need 50 samples of $f(\sigma)$ between 0 and σ_L. Figure 6-19 presents several spectra that have been actually computed from an interferogram using a sinc interpolation function limited to 10 sidelobes weighted or unweighted. Two cases are considered: weighted interferogram (giving an apodized ILS) or unweighted interferogram. The results show the improvement in interpolation accuracy due to weighting of the convolving function.

Altogether, computing 5×10^6 secondary samples from 10^6 primary ones takes 22 min on a 360/75 IBM. We can ask the question how much time has been saved compared with direct computation? We might have added to the actual 1024 K sample interferogram 3×1024 K zeros, and transformed the 4×1024 K pseudo-interferogram to get a spectrum with the same resolution, but 4 (not 5) times more samples. Supposing no limitation from disk capacity, the computation time would have been 36 minutes. The gain is even more spectacular if one uses the Array Processor, a special IBM unit in which the convolution operation is wired in. If it is connected to a 360/65 we have found the above cited 22 min time reduced 40 sec.

All the Aimé Cotton Laboratory spectra have been computed using this type of interpolation, at first from classically transformed primary samples and later using the FFT.

6-4 TREATMENT OF THE INTERFEROGRAM PRIOR TO FOURIER TRANSFORMATIONS

An actual interferogram is seldom identical to the ideally odd or even function so far considered. Let us remind the reader that while small imperfections in the optical surfaces of the interferometer merely give a slight reduction in S/N ratio, errors in the path difference measuring system distort the ILS.

We shall keep considering a minimally sampled interferogram (N samples) recorded by the stepping technique which guarantees an optimum S/N ratio from a given recording time T. Thus, we will not study the filtering or correcting techniques appropriate to a continuous scanning. We will also consider that the positioning accuracy of the samples is large enough to

make ILS distortions negligible.[4] The difficulties inherent in accurate measurement of intensities, with the large dynamic range necessary to accommodate simultaneously 10^6 spectral elements, can be considered solved and will not be treated here. We shall merely study two rather common causes of error in order to learn how to identify and eventually correct them. The first is a break in the sampling sequence and the second a shift in the samp' points (so called "phase error").

6-4.1 Break in the Sampling Sequence

Let us suppose the interferogram is a pure sine wave of wavenumber σ_0. If a sudden break in path difference occurs for $\delta = \delta_0$, the integral

$$f(\sigma) = \int_0^{\delta_{max}} \sin 2\pi\sigma_0\delta \sin 2\pi\sigma\delta \, d\delta,$$

will be replaced by

$$f'(\sigma) = \int_0^{\delta_0} \sin 2\pi\sigma_0\delta \sin 2\pi\sigma\delta \, d\delta + \int_{\delta_0}^{\delta_{max}} \sin 2\pi\sigma_0(\delta+\beta) \sin 2\pi\sigma\delta \, d\delta.$$

Thus, we can write

$$f'(\sigma) = f_1(\sigma) + f_2(\sigma) + f_3(\sigma),$$

$$f_1(\sigma) = \int_0^{\delta_0} \sin 2\pi\sigma_0\delta \sin 2\pi\sigma\delta \, d\delta = f(\sigma_0 - \sigma).$$

This is a normal ILS with resolution $\delta\sigma_1 = \frac{1}{2\delta_0}$.

$$f_2(\sigma) = \cos(2\pi\sigma_0\beta) \int_{\delta_0}^{\delta_{max}} \sin 2\pi\sigma_0\delta \sin 2\pi\sigma\delta \, d\delta$$

$$= \cos(2\pi\sigma_0\beta) f(\sigma_0 - \sigma),$$

$$\delta\sigma_2 = \frac{1}{2(\delta_{max} - \delta_0)}.$$

This is also a normal ILS, but multiplied by $\cos 2\pi\sigma_0\beta$ [see $h(\sigma)$ Figure 6-20]

$$f_3(\sigma) = \sin(2\pi\sigma_0\beta) \int_{\delta_0}^{\delta_{max}} \cos 2\pi\sigma_0\delta \sin 2\pi\sigma\delta \, d\delta$$

$$= \sin(2\pi\sigma_0\beta) h(\sigma_0 - \sigma).$$

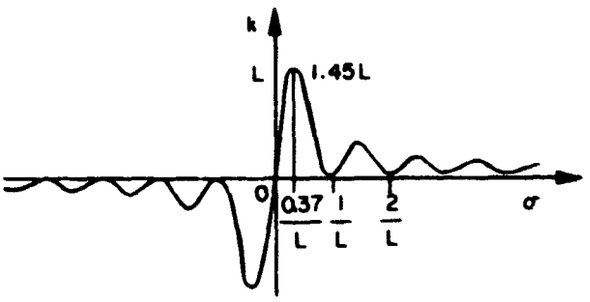


Figure 6-20. Sine Transform of a Square Pulse Function

The sum of these 3 functions can take different shapes. But in all cases there is both loss of resolution and production of oscillating secondary maxima which can extend very far from the central peak. Figure 6-21 gives a few examples of a distorted ILS (apodized).

The error is defined by two factors:

$$\alpha = \delta_0/\delta_{max},$$

$$r = \beta/\sigma_0.$$

Since the interferogram is weighted, it is normal that the highest distortions occur for small factors. In principle, the error can be corrected by programing provided the computed spectrum contains at least one well isolated line (emission or absorption) with a background which is known to be flat. However, we have not so far found it worthwhile to attempt the operation.

6-4.2 Phase Error

This has been much studied; a phase error can enter the interferogram under very different forms. So far we have considered an ideally odd or even interferogram recorded by an achromatic interferometer. In practice two different errors can take place.

(1) All the samples are shifted, none coincides with $\delta = 0$. This situation arises if the means for detecting zero path difference are not accurate enough.

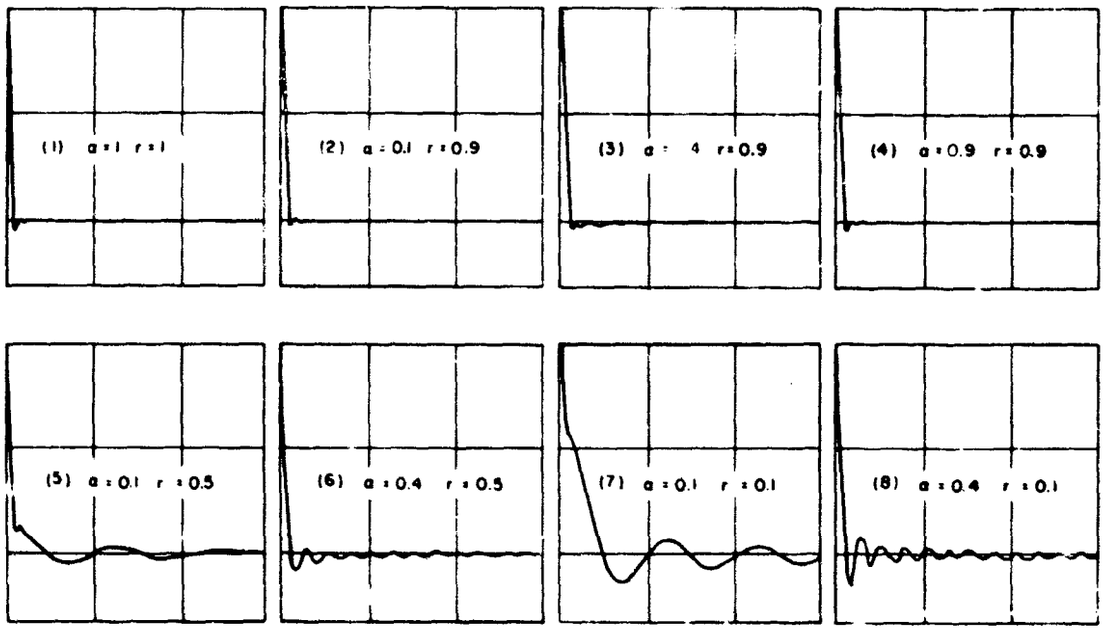


Figure 6-21. ILS After Break in the Sampling Sequence

(2) The interferometer is slightly chromatic; there is phase dispersion in the interferogram and the zero path difference point cannot be uniquely determined; the interferogram is neither odd nor even.

6-4.2.1 DISTORTION OF ILS ARISING FROM A CONSTANT PHASE ERROR

Let us consider an even interferogram (Figure 6-22*a*), but with a constant error in the path difference measuring device. When computing the spectrum we will in fact get the cosine transform of the even functions drawn on Figure 6-22*b* and *c*. What happens to the ILS?

$$F(\sigma) = \int_0^\infty A(\delta) \sin 2\pi\sigma_0(\delta+\epsilon) \sin 2\pi\sigma\delta \, d\delta$$

$$= \sin 2\pi\sigma_0\epsilon \int_0^\infty A(\delta) \cos 2\pi\sigma_0\delta \sin 2\pi\sigma\delta \, d\delta$$

$$+ \cos 2\pi\sigma_0\epsilon \int_0^\infty A(\delta) \sin 2\pi\sigma_0\delta \sin 2\pi\sigma\delta \, d\delta$$

$$F(\sigma) = \cos(2\pi\sigma_0\epsilon) f(\sigma) + \sin(2\pi\sigma_0\epsilon) h(\sigma).$$

This is the sum of 2 terms weighted by coefficients which depend only on the ratio ϵ/λ_0 of the path-difference error to the wavelength of the line of the required ILS. The first is the normal ILS, the second

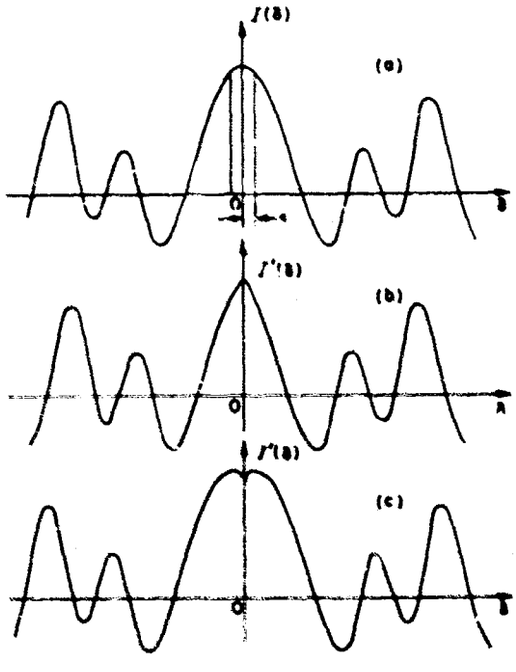


Figure 6-22. Effect of an Error in the Choice of Zero-Point on the Interferogram: (a) $\epsilon = 0$; (b) $\epsilon > 0$; (c) $\epsilon < 0$

the one we would get with $\epsilon = \lambda_0/4$. Figure 6-24 gives all the possible shapes the ILS will take for $0 < \epsilon/\lambda_0 < 0.25$ and Figure 6-25 for $-1 < \epsilon/\lambda_0 < +1$.

We must stress the extreme sensitivity of the ILS to very small shifts. We will show later that it is feasible to measure the position of lines in Fourier spectra down to a very small fraction of their width;

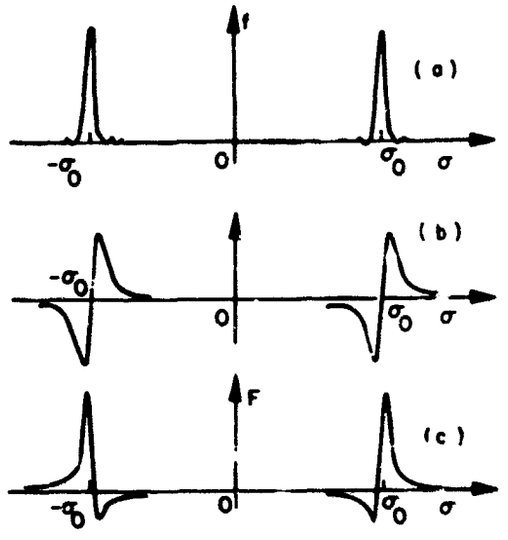


Figure 6-23. Apparatus Functions When There Is an Error in the Determination of $\delta = 0$

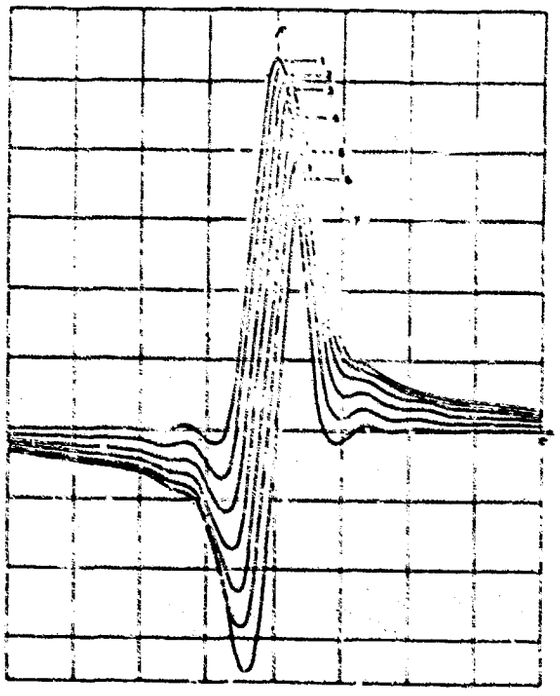


Figure 6-24. Different Forms of the Apparatus Function for Different Values of the Phase Displacement. Curve 1, $\epsilon/\lambda = 0$; curve 2, $\epsilon/\lambda = 0.042$; curve 3, $\epsilon/\lambda = 0.084$; curve 4, $\epsilon/\lambda = 0.126$; curve 5, $\epsilon/\lambda = 0.168$; curve 6, $\epsilon/\lambda = 0.210$; curve 7, $\epsilon/\lambda = 0.250$

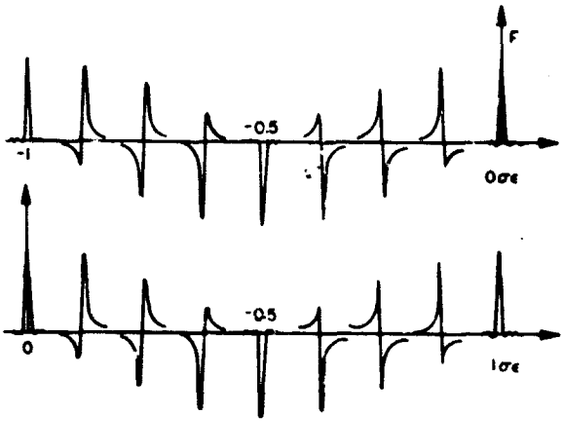


Figure 6-25. Deformation of the Apparatus Function for Phase Displacements Taken Between -2π and $+2\pi$

of course, for this to be meaningful the lines must be symmetric.

Furthermore, if we consider not just one line but an extended well-filled spectrum (either emission or absorption), the very small background level shifts will add systematically for all spectral elements and introduce an overall sizable shift. Figure 6-26 shows how an extremely small distortion of the ILS, corresponding to $\epsilon = \lambda/125$, can distort a rectangular spectrum the width of which is only 100 spectral elements. The effect increases with the number of spectral elements.[1]

6-4.2.3 RAPID REVIEW

The methods for eliminating the phase error from slightly chromatic interferometers are:

(1) If we record the interferogram from $-\delta_{max}$ to $+\delta_{max}$ and sum the squares of the sine and cosine FT, we can forget completely about phase. The method has the advantage of great instrumental simplicity. However, a double length interferogram, with twice as many samples, must be recorded. There is a degradation of the S/N ratio due to the use of nonlinear processes, and the zero intensity level is not as well defined.*

(2) Several different techniques have been proposed for improving asymmetrical interferograms of the type

$$I'(\delta) = \int_{-\sigma}^{+\sigma} B(\sigma) e^{-2\pi i \sigma \delta + \Phi(\sigma)}\, d\sigma$$

$$= \int_{-\sigma}^{+\sigma} B(\sigma) e^{-i\Phi(\sigma)} e^{-2\pi i \sigma \delta}\, d\sigma.$$

* All these diagrams illustrating the phase shift errors were obtained in 1953 on the 7094 IBM of Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

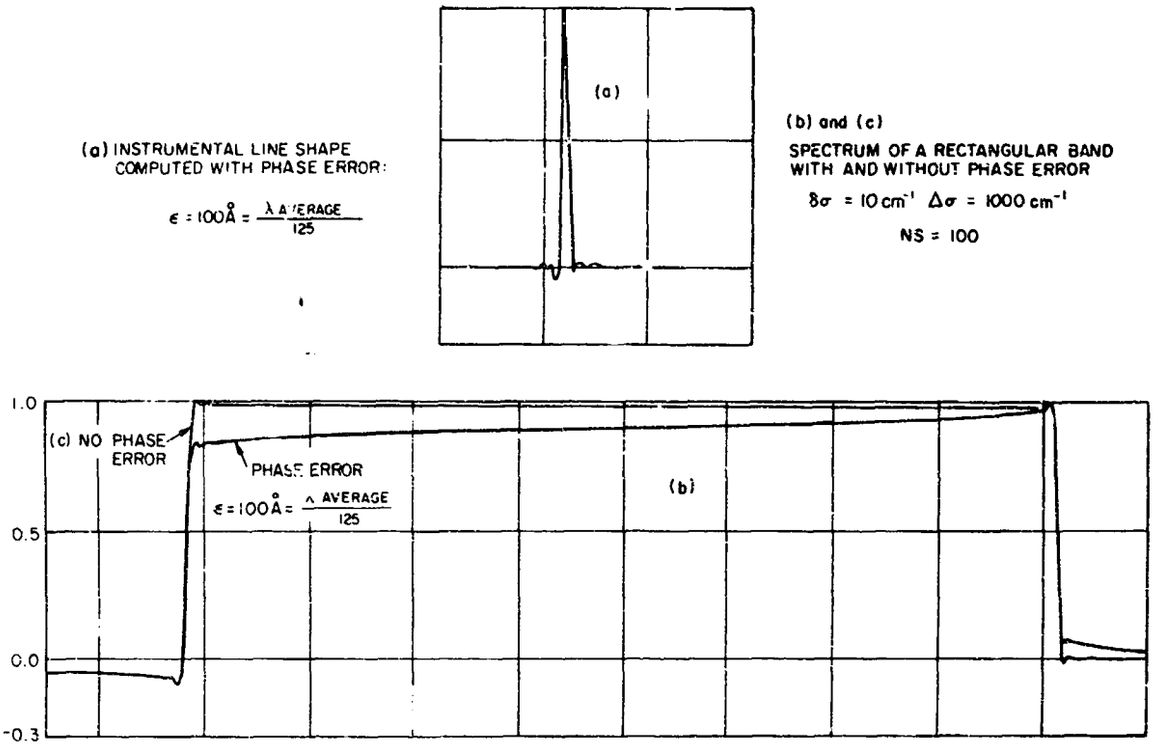


Figure 6-26. Effect of Distortion of the Instrumental Line Shape Computed With Phase Error $\epsilon = 100\ \text{Å} = \lambda_{\text{average}}/125$

with $\Phi(\sigma) = 2\pi\sigma\Phi'(\sigma)$. The FT of $I'(\delta)$ can be written as:

$$m(\sigma) = B(\sigma)e^{-i\Phi(\sigma)}$$

In order to reconstruct $B(\sigma)$ we must multiply $m(\sigma)$ by $e^{i\Phi(\sigma)}$. The whole difficulty lies in measuring $\Phi(\sigma)$. A really accurate determination would involve measuring the interferogram from $-\delta_{\max}$ to $+\delta_{\max}$. Accuracy can be improved by successive iterations; it is rather difficult to quote an exact figure.[25-27,30-32]

6-4.3 Zero Path Difference Error Correction

We aim at producing very high resolution ($\delta\sigma \simeq 5 \times 10^{-4}$ cm^{-1} corresponding to $R = 10^6$ at $\lambda = 2\mu$), so the wavenumbers must be measured with an accuracy of the order of 10^{-8}. This has been made possible by the construction of the 2 m path difference stepping interferometers.[4,27,38] While the accuracy in the sampling interval is always adequate, in some cases a recording technique is used in which no sample coincides with $\delta = 0$; the correction then must be introduced by the computer program.

6-4.3.1 ϵ IS KNOWN

The interferogram samples are taken at

$$\delta = \epsilon,\ \epsilon + h,\ \epsilon + 2h,\ \ldots,\ \epsilon + Nh.$$

The simplest correction procedure is the one of Loewenstein;[33-35] the spectrum is reconstructed by introducing the correction ϵ in the argument of the cosine.

One can easily show that the correction is complete only for the center of the ILS and poor for distant regions. Figure 6-27 gives the example of 2 lines at 7500 cm^{-1} and 8500 cm^{-1} corrected in this manner ($\epsilon = \lambda/9$). While the correction is adequate for the "true" spectrum, the sampling or aliases are not corrected, and they can be so poor that the true spectrum is distorted. Figure 6-28 presents the result in the case of a rectangular spectrum again with 100 spectral elements. Altogether, the method is adequate only for a narrow spectral range in the case where considerable oversampling is used because the images move very far apart.

However, this method does give the correct spectrum if the recording is made from $-\delta_{\max}$ to $+\delta_{\max}$.

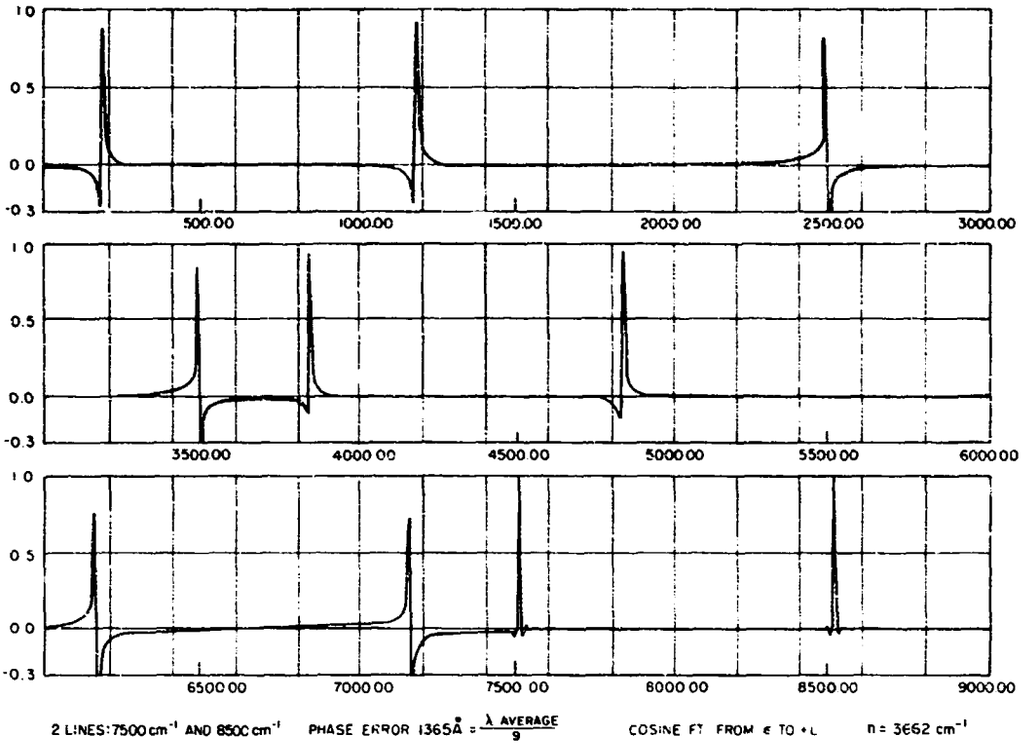


Figure 6-27. Two Line Spectrum Computed From ϵ to Δ_{max}

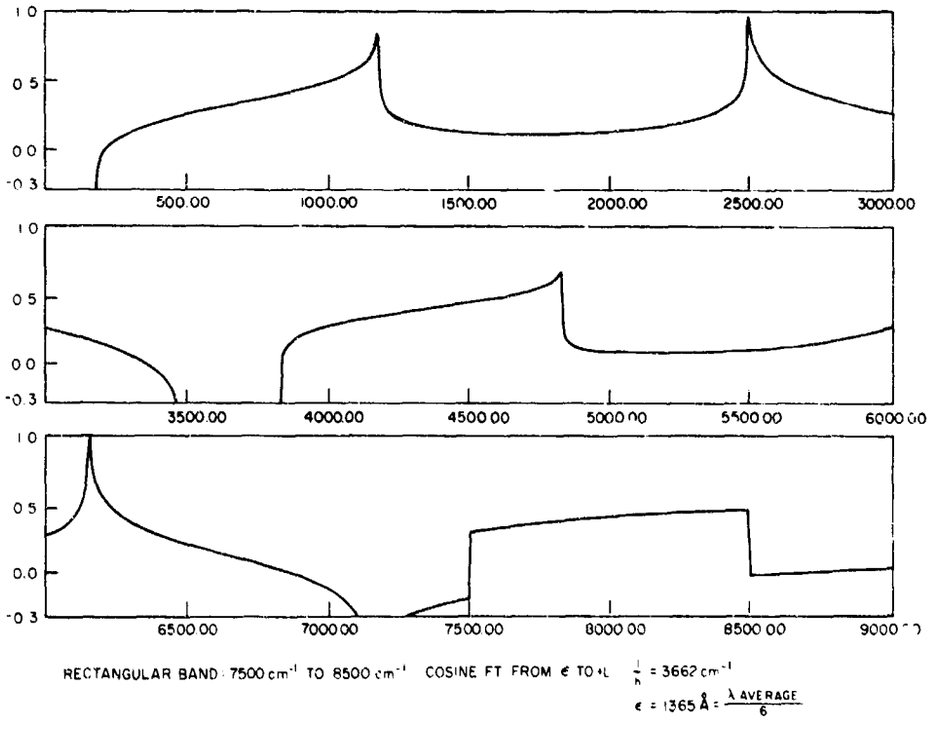


Figure 6-28. Continuum Spectrum Computed From ϵ to Δ_{max}

In this case:

$$B(\sigma) = \sum_{k=-N}^{k=N} I(kh+\epsilon) \cos 2\pi\sigma(kh+\epsilon),$$

or

$$B(\sigma) = \int_{-\delta_{max}}^{+\delta_{max}} I(\delta) R_h^\epsilon(\delta) \cos 2\pi\sigma(\delta+\epsilon)\, d\delta,$$

$R_h^\epsilon(\delta)$ being a Dirac comb with period h, shifted by ϵ with respect to the $\delta=0$ origin. One can see it as the convolution of the unshifted $R_h(\delta)$ comb by a shifted Dirac delta function $d(\epsilon)$. The FT $R_{1/h}^\epsilon(\sigma)$ of $R_h^\epsilon(\delta)$ is (within a factor $1/h$) the direct product of the normal $R_{1/h}(\sigma)$ Dirac comb by $\cos 2\pi\sigma\epsilon$ (Figure 6-29*a* and *b*). This means the true spectrum is reconstructed without any distortion; only the sampling images are multiplied by this coefficient. For instance, if we consider a line at $\sigma_0 = 8500\ \text{cm}^{-1}$ with $h = 0.27304\ 10^{-3}$ cm and $\epsilon = 0.136525\ 10^{-4}$ cm, we have the situation illustrated by Figure 6-30 in perfect agreement with expectations (Figure 6-29*c* and *d*). Figure 6-31 shows a rectangular spectrum in the same case.

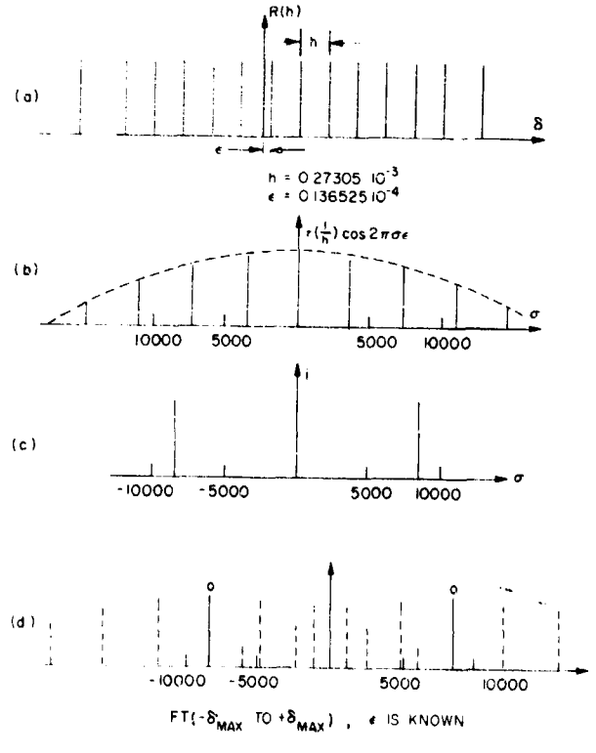


Figure 6-29. Computing With Phase Shift Error From $-\delta_{max}$ to $+\delta_{max}$

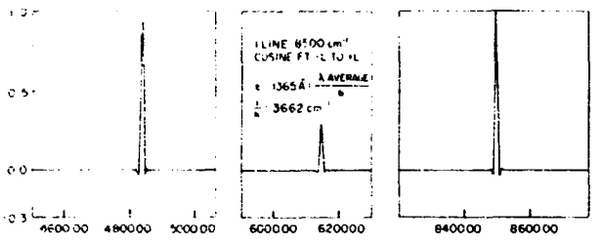


Figure 6-30. Line Spectrum Computed With Phase Shift Error From $-\delta_{max}$ to $+\delta_{max}$

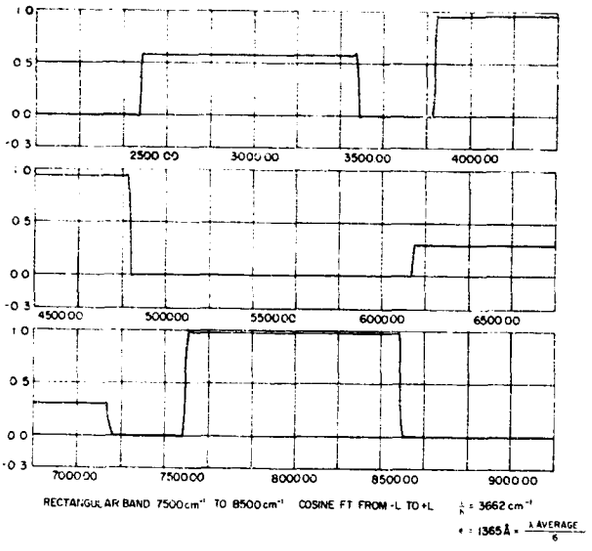


Figure 6-31. Continum Spectrum Computed with Phase Shift Error From $-\delta_{max}$ to $+\delta_{max}$

This technique is totally linear and does not introduce the difficulty of the power transform considered above, but a double length recording is still needed. Thus, we have used instead a rigorous interpolation procedure which, starting from the N actually recorded interferogram samples gives the set of N samples which would have been measured at $\delta = 0, h, 2h, \ldots, Nh$.

6-4.3.2 COMPUTATION OF AN INTERPOLATED "SECONDARY INTERFEROGRAM"

Let $I'(\delta)$ be the asymmetrically sampled interferogram (Figure 6-32*a*) and $B'(\sigma)$ the spectrum reconstructed by a [illegible] FT computed from $-\delta_{max}$ to

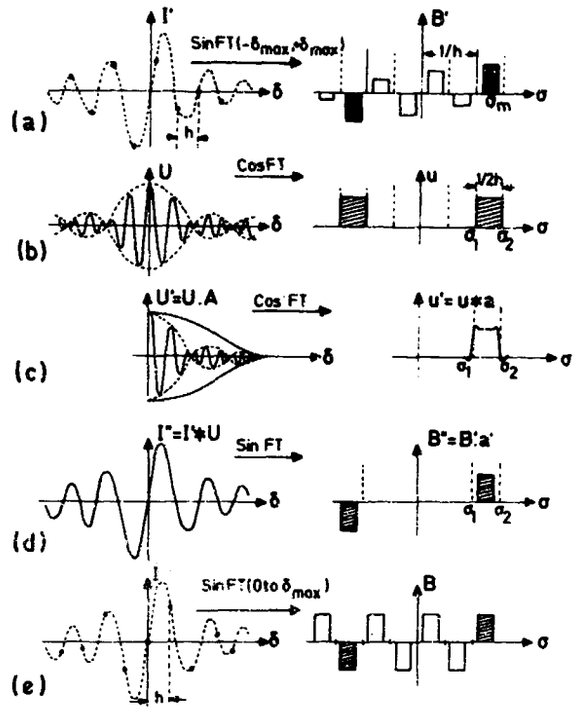


Figure 6-32. Interpolation in the Interferogram

$+\delta_{max}$. The original spectrum $B(\sigma)$ is surrounded by images, from which we can separate it by multiplying $B'(\sigma)$ by a rectangular "filtering" function $u(\sigma)$ of width $\frac{1}{2h}$, the FT of which is $U(\delta)$.

$$U(\delta)=\frac{\sin\frac{\pi\delta}{2h}}{\frac{\pi\delta}{2h}}\cos\frac{\pi}{2h}(4p+q)\delta, \qquad \text{(Figure 6-32}b\text{)}$$

where p is the number of ranges (of width $1/h$) which precede the one containing $B(\sigma)$; q is equal to 1 if the spectrum is in the first half of the range and to 3 if in the second half.

The interferogram $I''(\delta)$ which would be the FT of the $B'(\sigma)u(\sigma)$ spectrum is given by the convolution:

$$I''(\delta)=I'(\delta)*U(\delta). \qquad \text{(Figure 6-32}d\text{)}$$

In order to realize an ideal filtering this convolution should (again) stretch from $-\infty$ to $+\infty$; in practice it will be limited. Tests have shown that a highly accurate filtering can be realized using only a small fraction of $U(\delta)$ (of the order of 200 samples) provided suitable weighting is again used* (Figure 6-32*c*).

In this way any interferogram sample can be computed. In practice, N samples corresponding to $\delta=0,\ h,\ \ldots,\ Nh$ are computed for FT purposes (Figure 6-32*a*). In some cases we have computed samples with a separation much smaller than h, in order to trace a smooth interferogram curve with a plotter. Figures 6-33 to 6-39 show reconstructed interferograms corresponding to spectra of different widths; thus, the shape and the sampling intervals of the interferograms varied widely. The crosses mark the recorded points. The figures illustrate the sym-

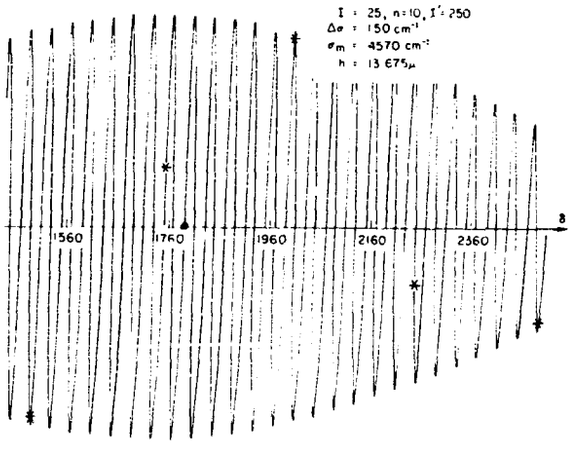


Figure 6-33. Example of Interpolated Interferogram

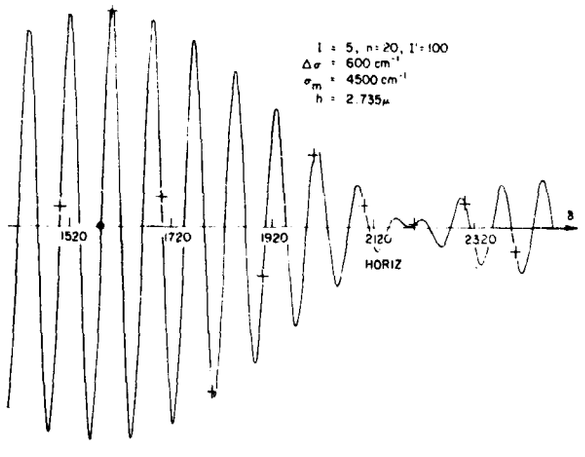


Figure 6-34. Example of Interpolated Interferogram

* Just as for the spectrum interpolation problem we use the weighting function $[1-\delta^2/(N\cdot h)^2]^2$.

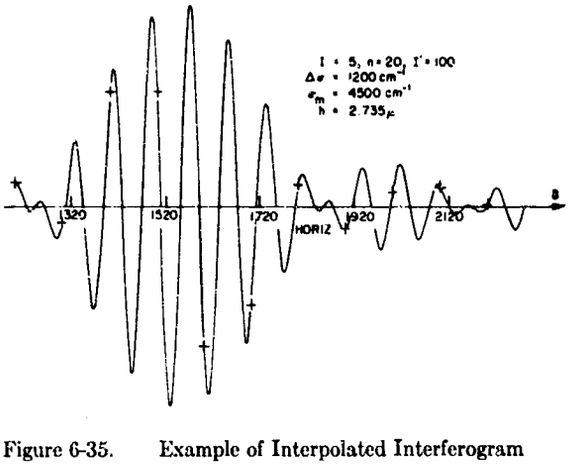


Figure 6-35. Example of Interpolated Interferogram

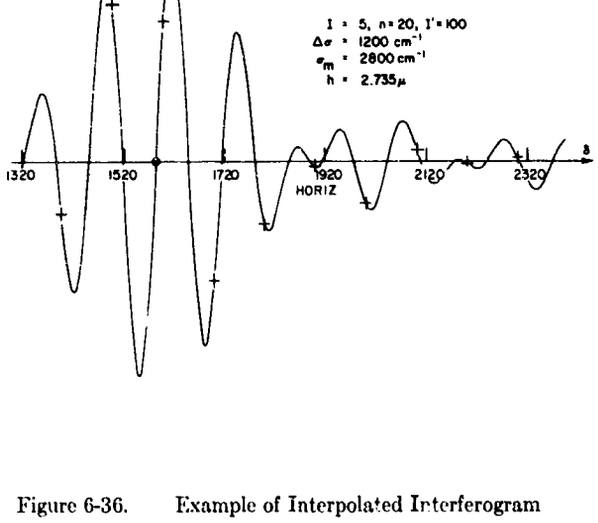


Figure 6-36. Example of Interpolated Interferogram

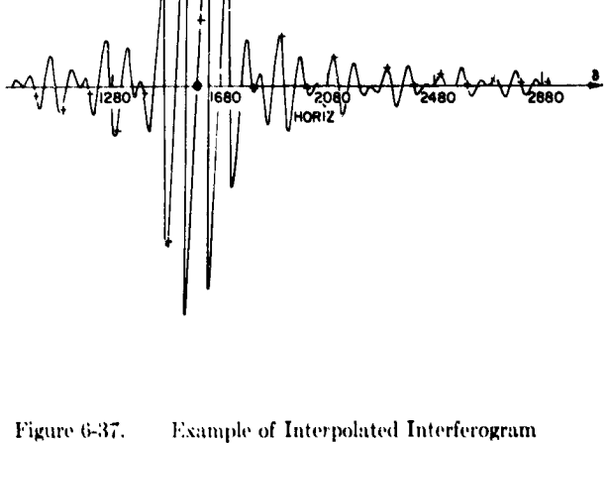


Figure 6-37. Example of Interpolated Interferogram

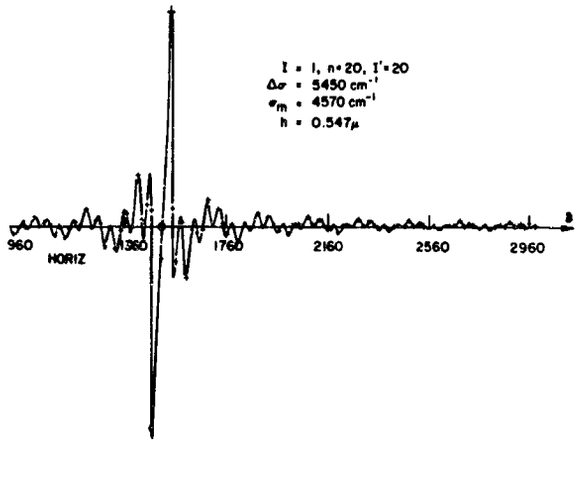


Figure 6-38. Example of Interpolated Interferogram

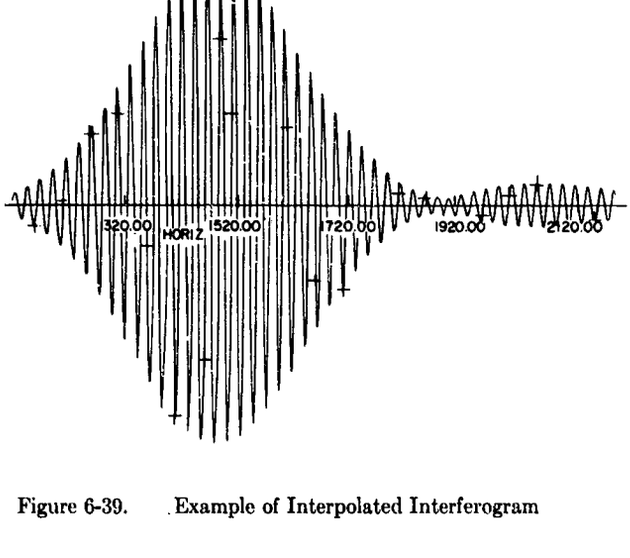


Figure 6-39. Example of Interpolated Interferogram

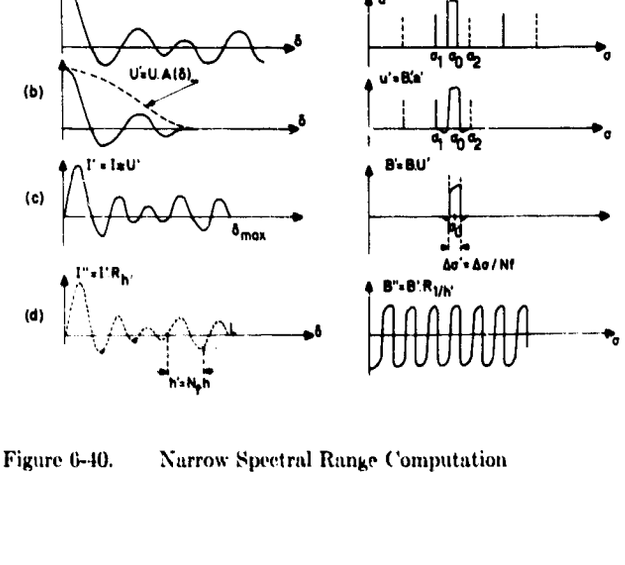


Figure 6-40. Narrow Spectral Range Computation

metry of the reconstructed interferograms. Figure 6-40 gives an example of a spectrum computed from an interpolated interferogram.

6-4.3.3 MEASURING THE PATH DIFFERENCE SHIFT ϵ

The shift ϵ can be measured rather simply from the complete FT of a small interferogram portion $I_p(\delta)$ approximately symmetrical with respect to zero path difference. We have

$$2\pi\sigma\epsilon = -\arctan\frac{b_e(\sigma)}{b_o(\sigma)}$$

with

$$b_e(\sigma) = \mathrm{FT}_{\cos} I_p(\delta)$$

$$b_o(\sigma) = \mathrm{FT}_{\sin} I_p(\delta).$$

Even if ϵ is computed for several σ and an average taken, the accuracy appears not quite good enough in some cases. A second "geometrical" technique has been used in the case of odd interferograms. Interferogram samples with a very small separation are computed close to the $\delta=0$ point. Then a linear interpolation is performed and the path difference giving zero intensity found.[31]

6-5 TEST PROGRAM

Even with the latest computers a full 10^6 sample FT, with subsequent interpolation and (especially) plotting, is a long operation. In a large computing center, with a great many customers it is somewhat difficult to have it done during daytime.

The ideal solution is of course real-time computation which is practical only with a special purpose computer directly linked to the experiment as described in Chapter 33.

However, *a posteriori* rapid computation is also very valuable. We have routinely used a program which computes 3 narrow spectral slices in about 5 minutes (one near the maximum expected signal), the others in regions of zero intensity; thus, both resolution and S/N can be estimated, and systematic errors detected. We have tried not only to reduce the computing time but also to simplify the data and program handling as much as possible. Only one run is needed to (1) read the data on magnetic tape, (2) compute, (3) plot the interpolated interferogram, (4) compute the full spectrum at low resolution, (5) compute and plot the 3 slices at high resolution, and (6) measure rms noise and S/N. Steps (2), (3), and (4) are often skipped.

6-5.1 Principle

Let us consider a narrow spectral range of mean wavenumber σ_0 (Figure 6-40*a*). We must isolate this range with a rectangular filter $u(\sigma)$, the FT of which is (Figure 6-40*b*)

$$U(\delta) = \frac{\sin \pi \Delta\sigma'\delta}{\pi \Delta\sigma'\delta} \cos 2\pi\sigma_m'\delta, \qquad (6\text{-}37)$$

where σ_m' is the central filter frequency, and $\Delta\sigma' = \Delta\sigma/N_f$ the width. The FT of the product $B(\sigma)\times u(\sigma)$ will be a "secondary" interferogram I''; it is the convolution of the recorded interferogram $I(\delta)$ by $U(\delta)$ (Figure 6-40*c*)

$$I''(\delta) = I(\delta) * U(\delta).$$

The proper sampling interval of $I''(\delta)$ will be (Figure 6-40*d*)

$$\frac{1}{2\,\Delta\sigma'} = h \times N_f.$$

Thus, starting from the primary interferogram with N samples we can obtain a secondary interferogram with only N/N_f points, the FT of which will give N/N_f spectral points. This means only a narrow spectral range will be computed but with the full resolution. Again the optimum compromise between computation time and accuracy of the results will be realized if the function $U(\delta)$ is weighted, which means using a filter with adequate shape.

6-5.2 Program

The program we have evolved takes the $N \leq 1024\ K$ primary interferogram and computes $8\ K$, $16\ K$, $32\ K$, $64\ K$ spectral slices from secondary interferograms with the same number of samples. The experimenter must provide the (approximate) central wavenumber of the wanted spectral slice and its width. Often several slices are needed; for instance, the one where peak signal is expected and another one with zero energy; thus, one can judge resolution and S/N ratio.

The computation time includes three terms corresponding to (1) production of the secondary interferogram by convolution, (2) FT, and (3) spectrum interpolation and plotting.

With the 360/75, computing slices of $8\ K$ each from 10^6 samples takes $3\times 96 = 288$ sec. With the already

mentioned Array Processor unit it is even possible to make the full spectrum computation by that process; 16 secondary interferograms of 64 K each are produced in 50 sec each; then 16 FFT operations are made (33 sec each). Altogether we need 18 min.

6-6 SOME OPERATIONS PERFORMED ON THE COMPUTED SPECTRA

A 10^6 samples spectrum contains so much information that the analysis of the spectrum has to be made automatic as far as possible.

6-6.1 Weighted Averaging of Spectra

The K identical resolution spectra computed from K identical length interferograms can easily be averaged. If they have the same S/N ratio, they are given equal weights and the resulting S/N is improved by a factor $\sqrt{K}$. This has been fully verified in many instances and shows the absence of systematic differences between successive spectra.

In the astronomical case, for reasons quite independent of the observers' will, many interferograms cannot be terminated, thus leading to lower resolution, but nevertheless usable spectra. They also have variable S/N ratios. In order to include them in the tracings given in the planetary Atlas we had to average spectra differing both in resolution and S/N ratio.

6-6.1.1 WEIGHTING COEFFICIENTS

If we have 3 spectra A, B, D, with the same resolution and S/N ratio $k_A > k_B > k_D$, the best spectrum A will be multiplied, before averaging, by $C_A = 1$ and the others by $C_B = (k_B/k_A)^2$ and $C_D = (k_D/k_A)^2$.

6-6.1.2 WEIGHTED AVERAGING

The solution we have used for variable resolution spectra involves making a weighted average of the primary (unapodized) spectral points of all available spectra, and then interpolating the result with a suitable function.

Suppose we have 4 spectra; the 4 interferograms (identically sampled) will be defined by a number of samples M_i and a weighting coefficient C_i (determined by analysis of the S/N ratio in the corresponding spectrum). In Figure 6-41 we have represented 4 rectangles of increasing lengths, proportional to M_i: $M_4 = M_1$; $M_3 < M_1 < M_2$.

In order to get the same sampling interval in the spectra the same number of samples must be used at the input of the FFT. Thus I_1, I_3, I_4 are first completed by adding $M_2 - M_1$, $M_2 - M_4$, and $M_2 - M_3$ zeros.

Each spectrum is individually normalized so that the peak intensity is unity; then the weighting coefficients are applied. Since the FT is a linear operation, the result is identical to the one which would have been realized from the FT of an interferogram $I'_m(\delta)$ weighted average of the individual interferograms. However, $I'_m(\delta)$ is not the wanted interferogram, since the samples do not have the same specific weight w.

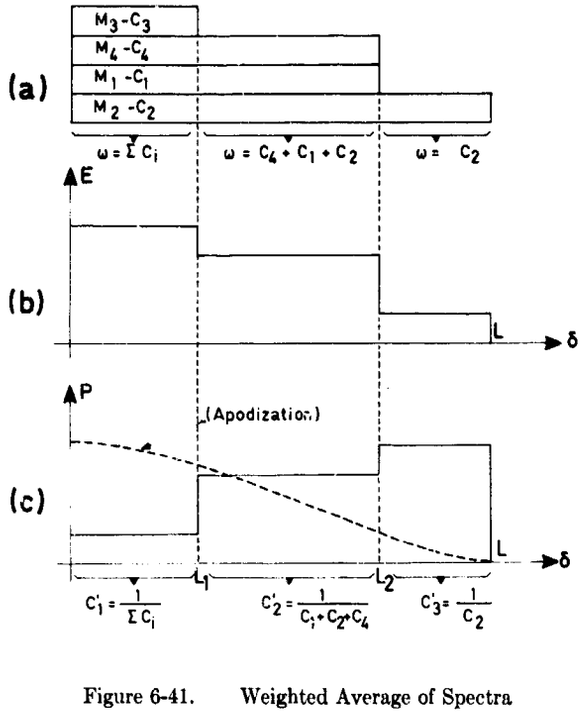


Figure 6-41. Weighted Average of Spectra

For samples

$$I \text{ to } M_3,\ w = C_1 + C_2 + C_3 + C_4;$$

$$M_3 \text{ to } M_4,\ w = C_1 + C_2 + C_4;$$

$$M_4 \text{ to } M_2,\ w = C_2.$$

In other words, the relationship between $I'_m(\delta)$ and $I_m(\delta)$ is:

$$I'_m(\delta) = I_m(\delta) \cdot E(\delta)$$

where $E(\delta)$ is a ramp function (Figure 6-41*b*). Thus:

$$I_m(\delta) = I'_m(\delta) \cdot \frac{1}{E(\delta)} = I'_m(\delta) \cdot P(\delta)$$

and

$$B_m(\sigma) = B'_m(\sigma) * FT[P(\delta)].$$

In order to find the correct spectrum $B_m(\sigma)$ we must convolve $B'_m(\sigma)$ by an interpolating function which is the FT of $I/E(\delta)$ (Figure 6-41c). Of course we can get an apodized spectrum, weighted by the $P(\delta)$ function by using the normal $A(\delta)$ apodization function.

The technique has been mainly used for the planetary spectra with $30000 < N < 60000$. The FFT were performed by adding zeros up to $64\,K$, and the primary spectra stored on magnetic tapes. The improvement in S/N ratio compared with the individual spectra was as expected.

6-6.2 Line Position Measurements

This must be done by the computer. The demonstrated accuracy in the high resolution spectra is $10^{-4}\,\mathrm{cm}^{-1}$. The spectral range given by a single spectrum is $9000\,\mathrm{cm}^{-1}$. To make full use of the available precision on a paper plot, supposing a pointing accuracy of about 1 mm, the full spectrum should be 90 km long.

It is also necessary to measure line intensity, equivalent width, and shape (for instance, asymmetry). All these measurements can be made from the secondary (interpolated) spectral samples stored on a magnetic tape. The tape also gives the overall number of spectral samples N_G, the wavenumber of the first sample, and the spectrum sampling interval $\delta\sigma_g$.

After a crude measurement on a paper plot the experimenter lists the approximate wavenumbers of the lines he is interested in. The computer searches for maximum and minimum intensity within a specified range (of the order of 2 to 5 times the resolution $\delta\sigma$, depending on spectral density, and so forth), then finds the centers of 10 chords at equidistant intensity (Figure 6-42 and 6-43). In general, the center of the chord closest to the line tip is taken as the line center.

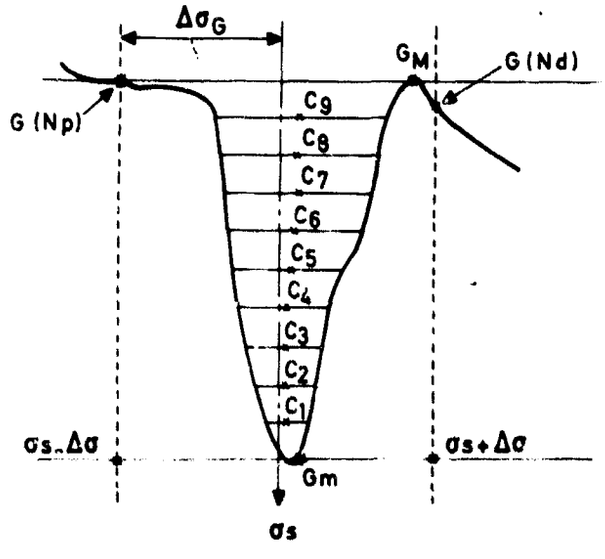


Figure 6-42. Line Position Measurement

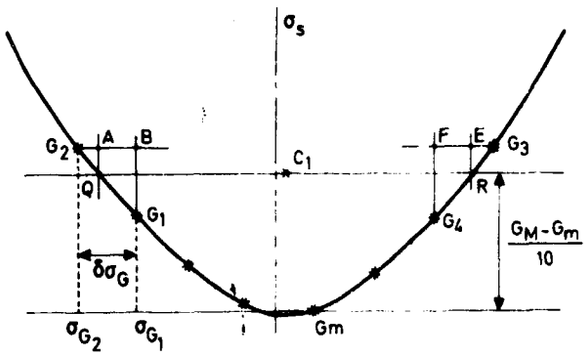


Figure 6-43. Line Position Measurement

Shifts between centers of other chords give line asymmetry.

Several corrections can be introduced at this stage. The first are experimental (reduction to vacuum, and correction for finite solid angle of the beams in the interferometer). These are applied simply by changing the reference wavenumber slightly. Doppler effect corrections on the astronomical spectra are also introduced at this stage.

6-6.2.1 ACCURACY

In order to test the accuracy of the entire program chain, we have performed the following operations on the 10^6 sample artificial interferogram corresponding to a line at $5000\,\mathrm{cm}^{-1}$, sampled with an $h = 4 \times 10^{-5}$ cm which means a $1/h = 12\,500\,\mathrm{cm}^{-1}$ in the computed spectrum; the resolution is $1.1 \times 10^{-2}\,\mathrm{cm}^{-1}$.

(1) Computed the 1024 K samples Fourier transform (the last locations being filled with zeros as explained above). The primary spectral sample separation is $\delta\sigma_B = 0.11920918955 \times 10^{-2}\,\mathrm{cm}^{-1}$.

(2) Computed the interpolated secondary spectral samples; their separation is $\delta\sigma_c = 0.2384185 79 \times 10^{-3}\,\mathrm{cm}^{-1}$.

(3) Computed the line position. The results are summarized in Table 6-2.

Table 6-2. Line Position

$2\,\Delta\sigma$ in cm^{-1}	Given wavenumber	Measured wavenumber
0.06	4999.9900	4999.9999997 86
0.11	5000.006	4999.9999997 86
0.20	5000.0500	4999.9999997 86
0.01	5000.0000	5000.0000000 76
0.02	5000.0000	5000.0000000 18
0.04	5000.0100	4999.9999997 86

6-6.3 Equivalent Widths

The measurement is performed within the range $\Delta\sigma_3$ on each side of the line axis as defined by the previous result. Figure 6-43 illustrates the absorption line case. Three parameters have to be found:

(1) The average background B_m; it is defined as the average of the highest intensities found to the right and left of the line within range $2\,\Delta\sigma_3$.

(2) the area S of the line below the background.

(3) the equivalent width (width of an equal area rectangle) (Figure 6-44).

The FFT and interpolations are performed in single precision since they provide the spectral intensities. The computation of the line positions uses double precision. We see the final error is of the order of 3×10^{-7} cm^{-1} or an accuracy of 1.6×10^{10}. Thus, computational errors are actually negligible; the best experimental results so far showing a 10^{-4} cm^{-1} error.

In order to make the full operation automatic we should not even ask the observer to detect the lines. Another program (not yet routinely used) has been written which computes the second derivative of the spectrum, thus eliminating the slow background level fluctuations. Figure 6-45 shows the result for a ICH_3 absorption band. Any ordinate above a given fraction, say 10 percent of the overall maximum, is taken as an indication of the presence of a line; the line position program is then used.

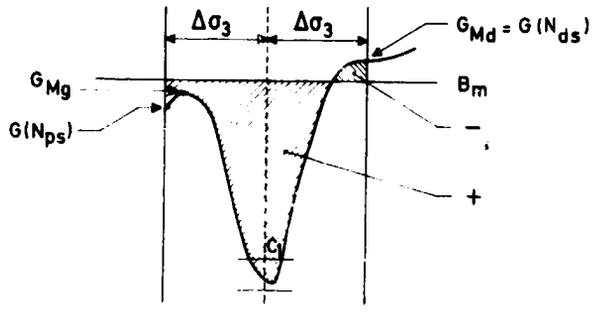


Figure 6-44. Equivalent Width Measurement

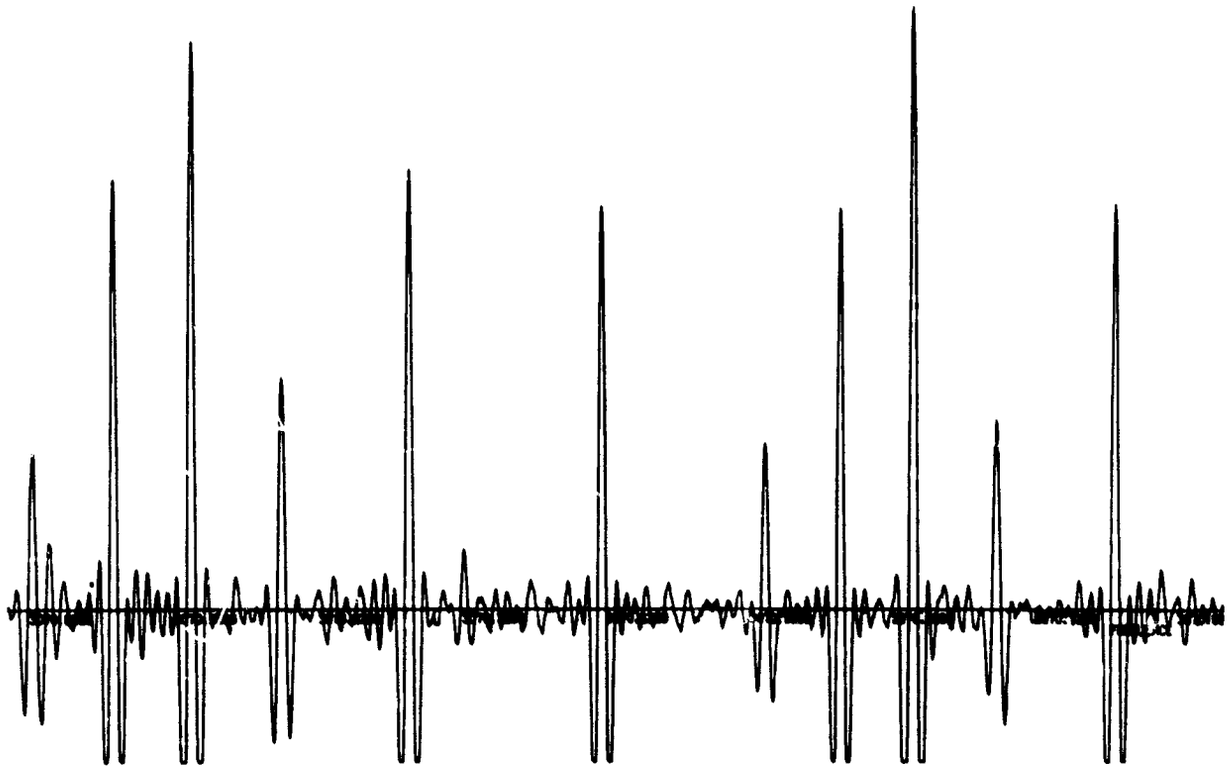


Figure 6-45. Second Derivative of an ICH_3 Spectrum

Appendix A

Hermitian Function

An Hermitian function is defined by

$$H(r)=H^*(-r).$$

It is a complex function with an even real part and an odd imaginary part.

$$H_r=A+jB,\ A \text{ is even, } B \text{ is odd.}$$

Its FT is real since

$$h(\delta)=\int_{-\infty}^{+\infty}(A+jB)e^{2\pi ir\delta}\,dr.$$

The real part is $\mathrm{FT}_{\cos}A-\mathrm{FT}_{\sin}B$.
The imaginary part is $\mathrm{FT}_{\sin}A+\mathrm{FT}_{\cos}B=0$.
Inversely, the FT of a real function is Hermitian.

Appendix B

FFT of an Even Real Function

In the case of an even real function I_k with $k\in(0, 1, \ldots, 2N-1)$

$$I_k=I_{2N-k},$$

$$I_{N+k}=I_{N-k}.\quad k\in(0, 1, \ldots, N-1).$$

We shall define G_k by

$$G_k=I_{2k}+iI_{2k+1}\quad k\in(0, 1, \ldots, N-1)$$

and we define the Hermitian function H_k by

$$H_k=\tfrac{1}{2}(G_k+G^*_{N-k}),$$

$$G^*_{N-k}=I_{2N-2k}-iI_{2N-2k+1}$$

$$=I_{2k}-iI_{2k-1},$$

$$H_k=I_{2k}+\frac{i}{2}(I_{2k+1}-I_{2k-1}).$$

Let the FT of H_k be called K_r.

$$K_r = \sum_{k=0}^{k=N-1} H_k W^{2kr}.$$

It is possible to compute the FT of an Hermitian function H_k with N complex terms from the N 2 complex terms of function F derived from H_k.

$$K_r = \sum_{k=0}^{k=N-1} I_{2k} W^{2rk} + \frac{i}{2} \sum_{k=0}^{k=N-1} (I_{2k+1} - I_{2k-1}),$$

$$= \sum_{k=0}^{k=N-1} I_{2k} W^{2rk}$$

$$+ \frac{i}{2} (1 - W^{2r}) \sum_{k=0}^{k=N-1} I_{2k+1} W^{2rk}.$$

$$K_{N-r} = \sum_{k=0}^{k=N-1} I_{2k} W^{2rk}$$

$$- \frac{i}{2} (1 - W^{2r}) \sum_{k=0}^{k=N-1} I_{2k+1} W^{2rk}.$$

On the other hand the DFT B_r of I_k can be written

$$B_r = \sum_{k=0}^{k=2N-1} I_k W^{rk},$$

$$B_r = \sum_{k=0}^{k=N-1} I_{2k} W^{2rk} + \sum_{k=0}^{k=N-1} I_{2k+1} W^{2rk} W^2.$$

We can now compute the N real samples of B_r

$$B_r = \tfrac{1}{2}(K_r + K_{N-r}) + \frac{K_r - K_{N-r}}{\sin \frac{\pi r}{N}}.$$

$$r(0, 1, \ldots, N-1)$$

References

1. Connes, J. (1961) *Rev. Opt.* **40**: 45.

2. Connes, J. (1961) *Rev. Opt.* **40**: 231.

3. Mertz, L. (1958) *J. Phys.* **8**: 233.

4. Connes, J., and Connes, P. (1966) *J. Opt. Soc. Am.* **56**: 896.

5. Shannon, C. E. (1949) *Proc. IRE* **37**: 10.

6. Shannon, C., and Weaver, W. (1949) *The Mathematical Theory of Communication*, Univ. Illinois Press., p. 53.

7. Kohlenberg, A. (1953) *J. Appl. Phys.* **24**: 1432.

8. Vanasse, G. A., and Sakai H. (1967) *Progress in Optics* **VI**: 260.

9. Connes, J. (1961) *Rev. Opt.* **40**: 171.

10. Cooley, J. W., Lewis, P. A. W., and Welch P. D. (1967) IBM research RC 1743.

11. Brenner, N. W. (1968) private communication.

12. Cooley, J. W., and Tukey, J. W. (1965) *Math. Comput.* **19**: 297.

13. *IEEE Transactions on Audio and Electroacoustics* (1965) **AU-15**: 56, the entire issue.

14. Gentlemen, W. M., and Sande, G. (1966) *AFIPS Proc.* **29**: 563.

15. Marshall C. Pease (1968) *J. ACM* **15**: 252.

16. Bergland, G. D., and Hale, H. W. (1967) *IEEE Trans. on Elect. Computers* **EC-16**: 180.

17. Bergland, G. D. (1968) *Math. Comput.* **22**: 275.

18. Bergland, G. D. (1967) *Math. Comput.* **21**: 236.

19. Singleton R. C. (1967) *Comm. ACM* **10**: 647.

20. Forman, M. L. (1966) *JOSA* **56**: 978.

21. Rigler, A. K. (1968) *JOSA* **58**: 274.

22. Brigham, E. O., and Morrow, R. E. (1967) *IEEE Spectrum* **63**.

23. Cochram, W. T., Cooley, J. W., and Farris, D. L., *et al.* (1967) *Proc. IEEE* **55**: 1664.

24. Bergland, G. D. (1968) *Comm. ACM* **11**: 703.

25. Connes, J., Connes, P., and Maillard, J. P. (1969) *Atlas des spectres dans le proche infra-rouge de Venus, Mars, Jupiter et Saturne*, Edition du C.N.R.S.

26. Buijs, H. L. (1969) *Appl. Opt.* **8**: 211.

27. Connes, J., Delouis, H., and Connes, P., *et al.* (1970) *New Rev. Opt. Appl.* **1**: 1.

28. Filler, A. S. (1964) *J. Opt. Soc. Am.* **54**: 762.

29. Papoulis, A. (1962) *The Fourier Integral and Its Applications* (McGraw Hill Book Company, Inc., New York) p. 51.

30. Mertz, L. (1965) *Transformations in Optics* (John Wiley and Sons, Inc., New York).

31. Forman, M. L., Steel, W. H., and Vanasse, G. A. (1966) *J. Opt. Soc. Am.* **56**: 59.

32. Sakai, H., Vanasse, G. A., and Forman, M. L. (1968) *J. Opt. Soc. Am.* **58**: 84.

33. Loewenstein, E. V. (1963) *Appl. Opt.* **2**: 491.

34. Mertz, L. (1963) *Appl. Opt.* **2**: 491.

35. Loewenstein, E. V. (1963) *Appl. Opt.* **2**: 1332.

36. Delouis, H. (1968) Thesis, Paris.

37. Connes, J., and Nozal, V. (1961) *J. Phys.* 92359.

38. Pinard, J. (1967) *J. Phys.* **28, C2**: 136.

Contents

7. The Impact of New Infrared Sources and Detectors on the Future of Fourier Spectroscopy

P. L. Richards
Department of Physics, University of California
and
Inorganic Materials Research Division,
Lawrence Radiation Laboratory
Berkeley, California

Abstract

Current developments in infrared sources and detectors are discussed with relation to their impact on the future of Fourier spectroscopy. References are given as a guide for further study.

7–1 INTRODUCTION

I believe that the pace of change in infrared techniques is increasing. Spectroscopists are becoming interested in new kinds of problems that place heavy demands on infrared apparatus. Partially because of these new problems, scientists from other disciplines are becoming increasingly interested in the infrared and are bringing their expertise and ideas. The interest in infrared astronomy is only one example of this trend. New developments in infrared sources and detectors on the horizon could radically change the experimental situation in the next few years. In these notes, I discuss several current developments that could prove important to the future of Fourier spectroscopy. References to review articles and to relevant new work are given as a guide for further study.

7–2 SOURCES

The now conventional sources of infrared radiation for laboratory spectroscopy include the glow-bar or Hg arc plus grating monochromator or Fourier interferometer. It is useful to think of both together as the source, because it is the continuous nature of the lamp emission that necessitates the use of the monochromator or interferometer. We may expect to replace both when a new source is found. This is certainly true when one of the many infrared gas lasers[1] is used as a spectroscopic source. Because of the lack of frequency tunability these lasers have had limited impact on infrared spectroscopy. Only in specialized problems, such as magnetic resonance in solids, where the effect under observation can be tuned, has the development of fixed-frequency infrared gas lasers been useful for this purpose. On the other hand, light scattering spectroscopy, which does not require source tunability, has prospered.

New types of laser sources are now being developed, using the techniques of non-linear optics, that provide tunable monochromatic radiation over a wide range of infrared frequencies. Continuous tunability is now available from 1 to beyond 250 cm^{-1} and over certain higher frequency bands. When fully developed, this class of sources could make the monochromator or Fourier spectrometer obsolete for laboratory infrared spectroscopy.

These new sources depend on the ability of a non-linear optical material to generate sum and difference frequencies, $\nu_3 = \nu_1 \pm \nu_2$, when two intense sources, ν_1 and ν_2, are present. In general, we consider three interacting waves with frequencies (energies) ν_1, ν_2, and ν_3, and k-vectors (momenta) $\boldsymbol{k}_1$, $\boldsymbol{k}_2$, and $\boldsymbol{k}_3$.

Both energy and momentum conservation are required. Momentum conservation depends on the index of refraction of the (anisotropic) non-linear crystal, so efficient non-linear processes will only occur for a specific relationship between frequencies, polarizations, and directions of propagation with respect to crystal axes.

Far infrared difference frequency generation was first demonstrated by Zernike and Berman.[2] They detected a difference frequency $\nu_3 \simeq 100\ \mathrm{cm}^{-1}$ when two frequency components, ν_1 and ν_2, of a broad band Nd glass laser were mixed in a quartz crystal. Recent experiments by Richards et al.[3] have found that the stability and power output of this system can be improved dramatically by mode-locking the Nd laser. Faries et al.[4] have demonstrated that continuously tunable difference frequency pulses with 10 W peak power can be obtained over the frequency range from 1 to $\sim 50\ \mathrm{cm}^{-1}$ by using two temperature-tuned ruby lasers as sources. Lithium niobate has replaced quartz as the non-linear crystal in most of these experiments.

Another difference frequency generation scheme, which is tunable in steps of $\sim 1\ \mathrm{cm}^{-1}$ from ~ 1 to $\sim 150\ \mathrm{cm}^{-1}$, uses two repetitively Q-switched CO_2 lasers operated on various rotational components of the lines around 9 to 10 μ. Semiconductors such as InSb or GaAs are used for non-linear mixers. Their cubic structure prevents the usual method of phase (momentum) matching by crystal orientation. Waveguide induced anisotropy, or modification of the free carrier contribution to the index by temperature or magnetic field, permits phase matching over part of the potentially useful frequency range.[5-7]

Yarborough et al.[8] have produced continuously tunable radiation from 50 to 150 cm^{-1} by a parametric procedure in $LiNbO_3$. A single-pulsed ruby laser with frequency ν_1 enters a crystal phase-matched to favor its splitting up into two beams, ν_2 and ν_3, that satisfy the requirements $\nu_1 = \nu_2 + \nu_3$ and $\boldsymbol{k}_1 = \boldsymbol{k}_2 + \boldsymbol{k}_3$. The low-frequency signal frequency, ν_3, generated is tuned by rotating the crystal.

We can see from the dates of referenced publications that tunable difference frequency generation has been developing very rapidly. The availability of organic dye lasers that can be tuned over the whole visible spectrum lets us speculate that continuously tunable difference frequency generators for the entire infrared range will soon be available. If such devices prove simple and reliable, they may replace presently used techniques (including Fourier spectroscopy) for laboratory absorption measurements.

7-3 DETECTORS

Improvements in detectors naturally affect both laboratory absorption spectroscopy and the measurement of emitting sources, whether in the laboratory, the atmosphere, or beyond. I expect this development to be fully as rapid as the progress of sources discussed above. The properties of the best available sources have a direct effect on the usefulness of Fourier spectroscopy. If the thruput requirements are not large, only the multiplex advantage favors Fourier spectroscopy over the diffraction grating. This advantage is realized only if the experiment is detector-noise limited. Developments in infrared detectors could therefore bring the simpler more robust grating instrument back to the forefront.

In order to illustrate some possibilities for infrared detector development, I will discuss first some potential improvements in the He temperature bolometer and then describe a new type of detector based on Josephson tunneling between two superconductors. These two topics by no means exhaust the possible developments, but serve to illustrate the growing importance of the techniques and materials of low-temperature physics on the detection of infrared radiation.

The most generally useful modern infrared detector is the Ge bolometer operated at pumped He temperature. These bolometers, first described by Low,[9] have been used by many workers. Such bolometers are often limited by room temperature background radiation noise in the near infrared (BLIP condition), but never in the far infrared. A recent development of great interest is the report by Drew and Sievers[10] of a Ge bolometer cooled by a He^3 refrigerator with a sink temperature of 0.37° K. The importance of lowering the heat-sink temperature below the usual 1.0 to 1.5° K obtainable from He^4 is illustrated by a calculation done by the author[11] that shows that if the bolometer parameters are optimized at each value of heat-sink temperature, the noise equivalent power varies approximately as T^3. This result is valid if the limiting noise is low-temperature Johnson noise or thermal fluctuation noise, but not, of course, if the limiting noise comes from the signal or background radiation.

Thus, a typical large ($\simeq 6\ \mathrm{mm}^3$) Ge bolometer for far infrared use with a sink temperature of 1.1° K and $\mathrm{NEP} \simeq 10^{-12}\ \mathrm{W}/\sqrt{\mathrm{Hz}}$ would correspond to $\simeq 3 \times 10^{-14}\ \mathrm{W}/\sqrt{\mathrm{Hz}}$ at 0.37° K (Drew and Sievers result) and $\sim 4 \times 10^{-18}\ \mathrm{W}/\sqrt{\mathrm{Hz}}$ at .02° K. A temperature of 20 mK or less can be reached by reliable (if troublesome) techniques of modern low-temperature physics, such as He^3-He^4 dilution refrigerators or adiabatic demagnetization. The NEP varies as the square root of the bolometer volume[11] so that our example of a 6-mm^3 bolometer for use in a far-infrared spectrometer has a limiting NEP at least a decade larger than a bolometer just large enough to see a diffraction-limited near-infrared image.

No one has constructed such a millidegree bolometer or even found the concentration of dopants in Ge required for one, but there is an excellent chance of success for anyone willing to take the trouble. The advantages in going to lower temperatures apply with small quantitative differences to any bolometer. It may be that they will be easier to realize with superconducting bolometers or other kinds. Certainly the superconducting bolometer has not yet been operated in a fully developed form.[12] The often quoted difficulty of regulating the temperature adequately to remain on the superconducting transition has been solved by modern low-temperature techniques.

A radically different type of detector that shows great promise is based on the supercurrents or Josephson currents that tunnel through a thin oxide barrier between two superconductors. This Josephson effect is one of the most complicated and interesting phenomena of modern solid state physics.[13] For the present discussion, we only point out that the current is the sine of the time integral of the voltage $I = I_0 \sin[\int^t 2eV(t')\,dt'/\hbar + \phi_0]$. This highly nonlinear current-voltage relation allows us to construct a number of useful devices. For example, if a dc voltage V_0 is applied, we have a quantum oscillator with frequency $\nu = 2eV_0/\hbar$, that is, 484 GHz per mV. Grimes, Richards, and Shapiro[14] showed that such a tunnel junction could be used to make a broad band far infrared detector with NEP$\simeq 5\times10^{-14}$ W/$\sqrt{\text{Hz}}$.

Richards and Sterling[15] have constructed a narrow band regenerative receiver in the far-infrared with peak NEP$\lesssim 10^{-14}$ W/$\sqrt{\text{Hz}}$. Grimes and Shapiro[16] have shown that Josephson junctions make good mixers. Since Josephson currents respond to rf voltages at frequencies as high as 10^{13} Hz, we can seriously consider constructing a heterodyne infrared detector with a bandwidth of several cm^{-1}. One of the tunable laser sources discussed above could be used for the local oscillator.

In their present form such detectors suffer seriously from a poor match to the incoming radiation. If the impedance match problem can be solved and comparable sensitivity obtained at high frequencies to that obtained for dc magnetic fields, NEP's in the 10^{-20} W/$\sqrt{\text{Hz}}$ region are possible. Such devices could count single microwave or infrared photons.[17,18]

In summary, it seems probable that developments in infrared technology, either in sources or detectors, will strongly affect the future usefulness of Fourier spectrometers. Infrared techniques are unlikely to remain as static in the next decade as they have in the past.

References

1. Patel, C. K. N. (1968) in *Lasers*, A. K. Levine, ed., Marcel Dekker, New York.

2. Zernike, F., and Berman, P. R. (1965) *Phys. Rev. Letters* **15**: 999.

3. Richards, Shelton, Shen, and Yang (1970) (to be published).

4. Faries, Gehring, Richards, and Shen (1969) *Phys. Rev.* **180**: 363.

5. Chang, T. Y., Van Tran, N., and Patel, C. K. N. (1968) *Appl. Phys. Letters* **13**: 357.

6. Zernike, F. (1969) *Phys. Rev. Letters* **22**: 931.

7. Van Tran, N., and Patel, C. K. N. (1969) *Phys. Rev. Letters* **22**: 463.

8. Yarborough, Sussman, Purhoff, Pantell, and Johnson (1969) *Appl. Phys. Letters* **15**: 102.

9. Low, F. J. (1961) *J. Opt. Soc. Am.* **51**: 1300.

10. Drew, H. D., and Sievers, A. (1969) *Appl. Optics* **8**: 2067.

11. Richards, P. L. (1968) *Far Infrared Detectors*, Proceedings of the NATO Advanced Summer School on Far Infrared Spectroscopy of Solids, Delft.

12. Putley, E., and Marlin, D. (1966) in *Millimeter and Submillimeter Waves*, D. Martin, Ed., North-Holland, Amsterdam.

13. Richards, P. L., Shapiro, S., and Grimes, C. C. (1968) *Am. J. Phy* **36**: 690.

14. Grimes, C. C., Richards, P. L., and Shapiro, S. (1966) *Phys. Rev. Letters* **17**: 431. Grimes, C. C., Richards, P. L., and Shapiro, S. (1968a) *J. Appl. Phys.* **39**: 3905.

15. Richards, P. L., and Sterling, S. A. (1969) *Appl. Phys. Letters* **15**: 394.

16. Grimes, C. C., and Shapiro, S. (1968b) *Phys. Rev.* **169**: 397.

17. Chiao, R. Y. (1970) *Bull. Am. Phys. Soc.* **15** (Ser. II): 39.

18. Richards, P. L. (1970) The Josephson Junction as a Detector of Microwave and Far Infrared Radiation in *Physics of III–V Compounds* Vol. 6, Academic Press, New York (to be published—reprints available from the author).

Contents

8. High Resolution and High Information Fourier Spectroscopy

Pierre Connes
Laboratoire Aimé Cotton
C.N.R.S. Orsay, France

Abstract

A summary of the Fourier spectroscopy work at Aimé Cotton Laboratory is presented. The original goal of computing 10^6 samples spectra for high resolution wide spectral range work in the near infrared has at last been reached; other problems, in particular planetary and stellar spectroscopy, have been tackled on the way.

8–1 INTRODUCTION

When Fourier spectroscopy work began at Laboratoire Aimé Cotton in 1954 under the direction of P. Jacquinot, the chief aim was the analysis of poorly understood atomic emission spectra. For most of the elements, the visible region had been extensively studied by photographic or photoelectric techniques; it was hoped Fourier spectroscopy would improve the results and extend them to the infrared.

From the very start, the laboratory had great confidence that high resolving power and high accuracy of wavenumber measurements could be obtained (after all, Michelson had done it already) and would prove the main attractions of the technique. How large a spectral domain could actually be treated in this way appeared quite unpredictable; and in order to better the Fabry Perot etalon and grating spectrometers both high resolution and wide spectral coverage were needed at the same time. This Michelson had not done!

Two classes of problems had to be solved; they can be listed under recording and computing.

8–2 INTERFEROMETERS AND RECORDING

One of the main advantages of Fourier interferometric spectroscopy is, of course, the multiplexing property, although other systems can multiplex in principle as well.[1] The theoretical gain relative to a scanning technique is easy to define; if M spectral elements are studied, recording time can be reduced by a factor M, or signal-to-noise ratio can be improved by $\sqrt{M}$, etc.[2] Elementary considerations show how the gain is much reduced in the photoemissive receiver range.[3] Unfortunately, non-fundamental factors should also be considered. In principle, the Fourier inversion process gives perfect discrimination between spectral elements, however large M is. In practice, interferogram distortions due to source intensity or recording system fluctuations, non-linearities, etc., produce random systematic spectrum distortions; these invariably increase with M.

A "quality factor" Q should be considered for each Fourier spectrum and plays as important a role as the multiplexing factor M. Q is simply the ratio between the total energy available in the spectral range and the minimum detectable energy in the spectrum, that is, the energy of a single line whose intensity is equal to the rms noise. One can consider either an emission or an absorption line, but its width must be negligible (smaller than the resolution). The noise must be the actually measured noise (from differences between successive spectra, or from the measured fluctuations in regions of known zero intensity), and not computed

from source energy and detector NEP for instance. Thus, all causes of error are included.

The Q factor can be written as

$$Q = M \frac{B_m}{N_{\text{eff}}}, \tag{8-1}$$

where B_m is the mean spectral energy over the free spectral range and N_{eff} is the rms noise. No simple considerations can be used to predict Q, although, of course, in each case photon noise would provide an upper limit.

The actually recorded Q depends first on source fluctuations. Easy problems include (for instance) laboratory absorption spectra; difficult cases include many emission sources and all astronomical ones, because of turbulence. However, suitable recording techniques can greatly improve Q for a given source; this is the whole story of astronomical Fourier spectroscopy. The factor M indicates the gain to be realized in principle from multiplexing; but unless enough care is used, Q will not be increased sufficiently and the results will be uselessly spoiled.

In short, M measures the profit and Q the labor. We had no inkling in those early days of the labor involved, or we would have turned to easier problems. The main steps shall be briefly recalled. On the way to the original goal, altogether different problems were tackled first, and techniques initially developed for one reason were kept for another.

The first interferometer involved no new features. It was used on night sky emission;[4] the interferograms were pen recorded and then sampled by hand and eye. The maximum was $N = 6000$ input samples and resolving power was 2000. This problem was excellent for beginners; it was one in which a large gain could easily be realized by elementary techniques, simply because the results given by grating spectrometers were so poor ($R = 150$ at 1.6μ).

The three main features of the interferometers developed during the next years were cat's eyes retroreflectors, step-by-step recording with interferometric servo control, and internal modulation.

Cat's eyes retroreflectors[6,7] were originally considered the easiest solution to the misadjustment problem. Today, with the experience gained in servo controlling path difference, a classical plane mirror Michelson with parallelism servo control added is preferred. However, the small mirror in the cat's eye can easily be supported by a piezoelectic ceramic and provides a very convenient way of increasing the path difference servo speed and of producing internal modulation.[10,12]

Step-by-step recording with interferometric servo control was initially developed because of the economy of sampling points obtained in this way; continuous scan techniques need oversampling to obtain the optimum signal-to-noise ratio.[5] When the astronomical interferometer was built (1963) the largest number of samples $N_{\max}$ that could be transformed in practice was of the order of 10^4; if oversampling had been used, there would have been no hope of exceeding grating spectrometers (in the case of bright stars or planets) since they already gave resolving powers of the order of $2 \cdot 10^3$.

Huge improvements in computing Fourier transforms have taken place since; this means it has become quite easy to oversample in most cases. One would not bother today to reduce the number of samples to the minimum sampling theorem figure for N less than a few times 10^5; nevertheless, it is still necessary for $N > 10^6$, which corresponds to the present state of the art.

However the stepping-recording technique has in the meantime proved of unequalled accuracy in instrumental line shape definition, spectrum reproducibility[6,7] and wavenumber precision.[10] The possibility of interrupting or slowing down the recording has proved invaluable, especially in the astronomical case.[6,8,11]

The internal modulation technique was imagined by Mertz[13] but was dropped in favor of "fast scanning" with later "coherent adding."[14,15] While well suited to the low N problems, fast scanning could be extended to high N only with disk or drum storage devices and would lose much of its original simplicity.

Internal modulation was developed only when the laboratory became interested in astronomical applications,[6] since the improvement is large mainly when a fluctuating source with a more or less $1/f$ spectrum is observed. However, it proved so convenient that it is now used in all cases.[10,12] With internal modulation, recording time is unlimited and no adding of the interferograms is required. The modulation frequency which best suits the detector or the problem can be selected at will.

A last remark on a generally misunderstood point should be added here. The so-called dynamic range problem in Fourier spectroscopy does not exist and the simple and obvious technique of changing the gain after the signal amplitude has dropped[8,12] makes it possible to cope with the worst cases. It is equivalent to using coarser quantization levels for the central peak; the errors are found totally negligible in practice.

8-3 COMPUTING

We were at first very slow to understand the value of purely digital Fourier transformation, and turned to general purpose computers only after Gebbie presented his first results[16] at the 1957 Bellevue conference. However, computers seemed usable only for transforming a few thousand samples; so continued work on a special purpose hybrid computer proved to be a "blind alley." This seemed the only way of reaching our goal, the transformation of 10^6 samples. The instrument should have provided the result in the incredibly short time of just a few days! (The latest

1970 digital computer figure is 9 min.) The electronic part was actually built and worked quite well;[17] however, operation implied the rather naïve assumption that a magnetic tape loop could be rotated perhaps 10^5 times without wear and the actual results were a great shock. When this stage was reached, small magnetic drums were becoming available, but purely digital computation techniques had progressed with giant strides and the system was abandoned.

The successive steps in the development of programming general purpose computers were described by J. Connes in Chapter 6 (see Figure 8-1). Each step

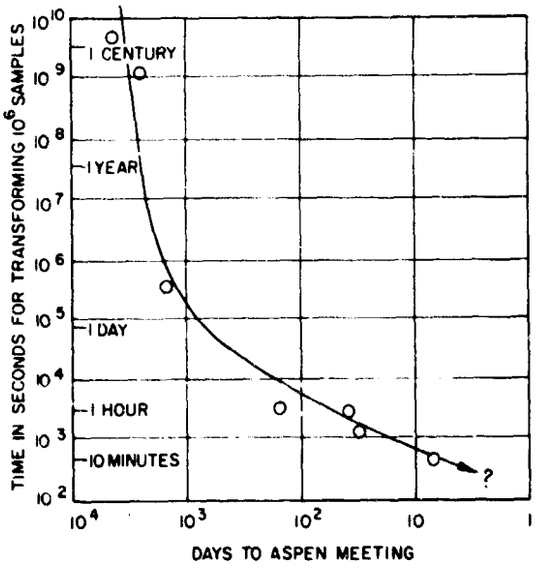


Figure 8-1. Time in Seconds for Transforming 10^6 Samples as a Function of Days to the Aspen Meeting

was eagerly awaited and immediately put to use. During the first astronomical observations of 1966, a large number of rolls of paper tape of up to 60,000 samples piled up with no means of computing the complete spectra (consisting of 60,000 samples). The extension of the Cooley-Tukey fast Fourier transform method to large numbers of samples (greater than the central storage capacity) saved the situation shortly after the observations.

Fast access to the computer has always been a major wish. A telephone line transmission was used during the astronomical observations;[8] computation of a spectrum slice at maximum resolution was obtained about an hour after the end of the recording. This seemed wonderfully fast, but still most of the time was spent handling paper or magnetic tapes, or fighting transmission errors and not in computing. "False" real time computation was also used: a low resolution spectrum was computed from the initial portion of the interferogram and the results were transmitted to the observer before the end of the recording.

This technique was largely responsible for the success of these observations and for a time this appeared to be the way of the future. However it gets far too expensive when a large number of samples must be transmitted—unless one has the rare luck of operating close to a willing large computer.

Small general purpose computers have since become common and have been successfully programmed for real time operation (Bras et al. at the Aspen Conference). They cannot, however, handle our large amount of data; more precisely they are not fast enough for high input data rates. Now, again, the ultimate solution appears to have been found. The integrated circuits revolution has opened entirely new possibilities, and even a laboratory which is not basically computer oriented can now build its own. The real time classical Fourier transform computer was constructed by G. Michel who will give the description (Chapter 33). At long last the virtues of digital techniques, and this is a completely digital system, are thoroughly convincing. The main features are: 1) a basic computation time of 10^{-6} sec/input point/output point, which makes it about 10^3 times faster than a "small" general purpose computer programmed for the same problem; and 2) a maximum output capacity of 20,000 spectral samples (when complete), with no limitation on input capacity.

This computer is used for real time checking of a spectrum slice; both extended or simplified versions could be built for problems differing from ours.

8-4 RESULTS

The greater part of the effort has been, rather unexpectedly, devoted to astronomical sources. The first interferometer built and used in collaboration with the Jet Propulsion Laboratory, featured 5 cm carriage displacement (0.1 cm^{-1} resolution) and a maximum sampling speed of 5 samples/sec, which meant that 60,000 samples could be recorded in about 3 hours. This turned out to be the practical limit, partly because of the computing facilities. The spectral range which can be computed from one interferogram is then 3000 cm^{-1} when working at 0.1 cm^{-1} resolution.

An improvement of at least 100 in resolving power was realized on Venus, Mars, and Jupiter compared to the best grating spectra;[8,27] an elementary calculation shows that grating spectrometers would have required an increase of 10^{10} in recording time to do the same. Scientific results include the detection of HF, HCl, and CO on Venus and CO on Mars,[19,20] improved knowledge of their atmospheres [22,23,24] and, also unexpectedly, of the CO_2 molecule itself.[25] The complete spectra of Venus, Mars, Jupiter, and Saturn have been published in atlas form[26] (Figure 8-2). The spectra of a few bright cool stars have also been studied, and the technique appears powerful for the study of stellar atmospheres[28,29,30,31] and

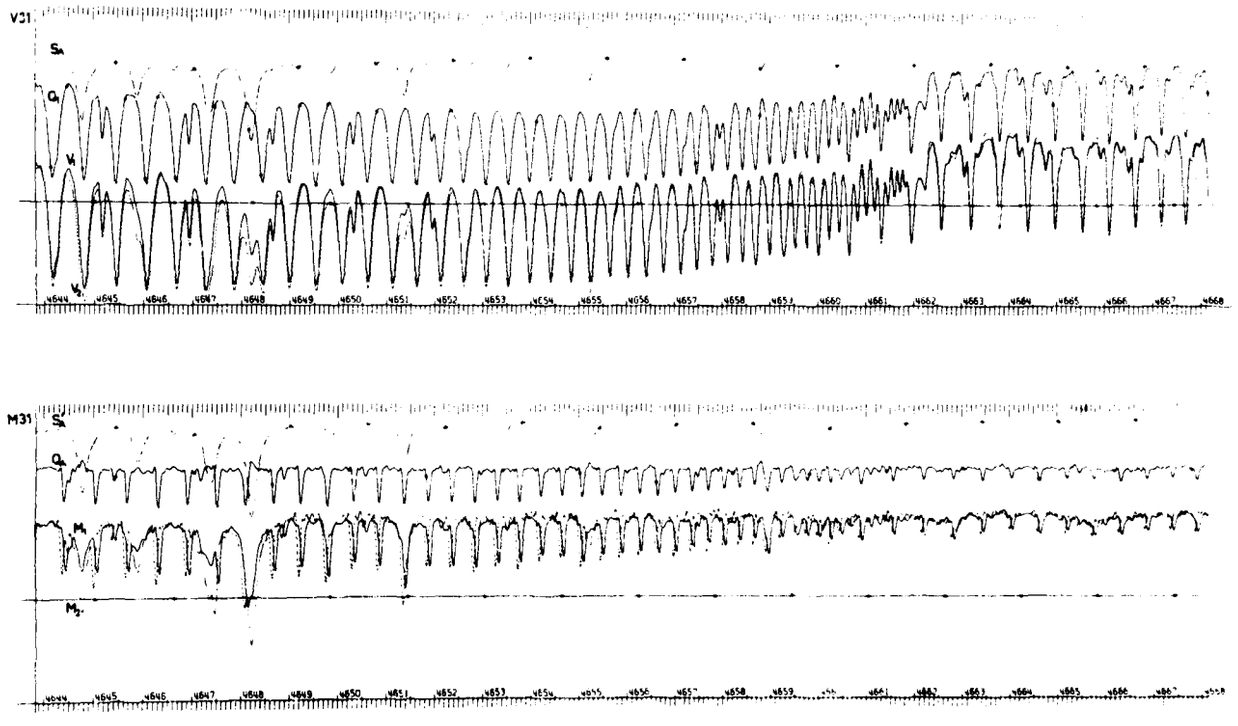


Figure 8-2. One Page of Near Infrared Planetary Atlas[26]

radial velocities.[32] In Figure 8-2, the scale is in vacuum wavenumber; line position measurements can be done from the plates to 10^{-2} cm^{-1} and from the digital data to 10^{-3} cm^{-1}. In the upper part, S_A is the solar spectrum, and V_1, V_2 are Venus spectra recorded at different elevations. Differences between V_1 and V_2 indicate telluric lines. In Q_1, the ratio of V_1/S_A, telluric absorption are nearly cancelled. In the lower part, S'_A is the solar spectrum, and M_1 and M_2 are Mars spectra recorded at different elevations and also with different planetary Doppler shifts. Q_A is the ratio M_1/S'_A. All planetary lines on the page are due to CO_2 bands; their intensity is much greater on Venus.

The second generation interferometer has been built by J. Pinard.[9,10] Maximum path difference has been increased by a factor of 20 to nearly 200 cm; the resolution is thus $5{\cdot}10^{-3}$ cm^{-1}. Since the speed of operation and maximum number of samples were about the same as for the first instrument, the spectral range covered from one interferogram was 20 times smaller, or about 150 cm^{-1}. This is too small for the original line emission problem; the instrument was mostly used on laboratory absorption spectra. A sizable gain in resolving power was realized compared to the best grating spectrometer results; furthermore, all the techniques used in this interferometer could readily be extended to larger path differences if the need arose. Thus, instruments of this type represent a final step in the spectroscopist's long quest for ever increasing resolving power.

The gain in wavenumber accuracy is even greater (1.2×10^{-4} rms error vs 2.1×10^{-3}). For lack of a long path cell, studies have been restricted to a few very absorbing gases (for instance CO, N_2O, ICH_3). The spectra are analysed by the group of Professor H. Amat (Faculté des Sciences de Paris).

The recently completed third generation interferometer[12] is described by G. Guelachvili and J. P. Maillard in Chapter 12. Since there appeared no pressing need of increasing the resolving power, the same maximum path difference was kept; however, the speed of the stepping drive was increased; up to 200 samples/sec could be recorded while keeping the internal modulation (the present stepping recorder gives a 50 samples/sec limit). For sources intense enough to justify still higher speeds, internal modulation would no longer be useful and the system would operate in the fast scanning mode (but without subsequent adding of interferograms).

Other new features include larger optics (8 cm diameter beam size) and a predominantly digital servo system giving much smaller minimum stepping interval (350 Å) and more freedom in the choice of free spectral range and amplitude of modulation.

The instrument has been used with up to 10^6 samples, which is not supposed to be any kind of limit. This applies again to minimally sampled interferograms, that is, the samples are all independent and actually produce 10^6 meaningful spectral samples. With a free spectral range extending from 0 to 11,000 cm^{-1}, the resolution is $\simeq 0{\cdot}10^{-2}$ cm^{-1} (apodized).

The first tests on rare earth emission spectra have shown a quality factor Q of $3 \cdot 10^7$. Since the sources are particularly troublesome and much work could be done to improve them, this present figure is again not a limit.

In this manner, the original program has at long last been fullfilled. From a single recording lasting a few hours it is now possible to obtain essentially all the information present in the spectrum. More precisely, hyperfine and isotopic structure, Zeeman patterns and line width and shape can be measured across the entire infrared range; the only limitations are due to source transmission and receiver characteristics.

A program of this type would have previously involved hundreds of different Fabry Perot or SISAM[33] recordings. In fact, it has never been done and analysis was always limited to a small number of bright lines.

8–5 CONCLUSION

Only technical problems have been solved, and this large amount of information is of course only what could be expected from properly used multiplex Fourier interferometric spectroscopy.[1] How much of the experience will be valuable to others is not clear; the case is a rather extreme one and most Fourier spectroscopists can get the results they are interested in using simpler interferometers and recording or computing schemes. However, we hope to have shown that the advantages of high resolving power and high accuracy of wavenumber measurement across broad spectral ranges are actual facts. As a consequence, and contrary to a rather common belief, Fourier spectroscopy is here to stay irrespective of any advance in the infrared detector art, and, on the contrary, an extension of the technique to short wavelengths is a most likely development.

References

1. Fellgett, P. B. (1967) *J. Phys.* **28**, C2: 165.
2. Jacquinot, P. (1958) *J. Phys.* **19**: 223.
3. Kahn, F. D. (1959) *Astrophys. J.* **129**: 518.
4. Connes, J., and Gush, H. (1960) *J. Phys.* **21**: 645.
5. Connes, J. (1961) *Rec. Opt.* **40**: 45, 115, 171, 185, 231.
6. Connes, J., and Connes, P. (1966) *J. Opt. Soc. Am.* **56**: 896.
7. Cuisenier, M., and Pinard, J. (1967) *J. Phys.* **28**, C2: 97.
8. Connes, J., Connes, P., and Maillard, J. P. (1967) *J. Phys.* **28**, C2: 136.
9. Pinard, J. (1967) *J. Phys.* **28**, C2: 136.
10. Pinard, J. (1970) *Ann. Phys.* 2.
11. Maillard, J. P. (1967) Thesis, University of Paris.
12. Connes, J., Delouis, H., Connes, P., Guelachvili, G., Maillard, J. P., and Michel, G. (1970) *Nouv. Rev. Opt. App.* 1: 3.
13. Mertz, L. (1958) *J. Phys.* **19**: 233.
14. Mertz, L. (1965) *Transformations in Optics*, Wiley, New York.
15. Mertz, L. (1965) *Astron. J.* **70**: 548.
16. Gebbie, H. A. (1958) *J. Phys.* **19**: 230.
17. Connes, P., and Michel, G. (1967) *Rev. Phys. Ap.* **2**: 175.
18. Connes, J., and Connes, P. (1967) *J. Phys.* **28**, C2: 37.
19. Benedict, W., Connes, J., Connes, P., and Kaplan, L. D. (1967) *Astrophys. J.* **147**: 230.
20. Benedict, W., Connes, J., Connes, P., and Kaplan, L. D. (1968) *Astrophys. J.* **152**: 731.
21. Kaplan, L. D., Connes, J., and Connes, P. (1967) *Astrophys. J.* **157**: L.187.
22. Young, L. D. J. (1969) *Icarus* **2**: 66.
23. Welch, W. J., and Rea, D. G. (1967) *Astrophys. J.* **148**: L.151.
24. Swings, P. (1969) *Proc. Am. Phil. Soc.* **113**: 229.
25. Cihla, Z., and Chedin, A. *J. Mol. Spectry.* (to be published).
26. Connes, J., Connes, P., and Maillard, J. P. (1969) *Atlas des Spectres Planétaires Infrarouges*, Editions du C.N.R.S., Paris.
27. Connes, J., and Connes, P. (1968) *Infrared Astronomy*, Gordon and Breach, New York, p. 193.
28. Connes, P., Connes, J., Benigue, R., Querci, M., and Chauville J. (1968) *Ann. Astrophys.* **31**: 485.
29. Montgomery, E. F., Connes, P., Connes, J., and Edmonds, F. N. (1969) *Bull. Am. Astr. Soc.* 1: 201.
30. Montgomery, E. F., Connes, P., Connes, J., and Edmonds, F. N. (1969) *Astrophys. J. Suppl.* **19**: 1.
31. Spinrad, H., and Wing, R. F. (1969) *Ann. Rev. Astr. Astrophys.* **7**: 249.
32. Bopp, B. W., and Edmonds, F. N. (1970) *Pub. Ast. Soc. Pac.*, **82**: 299.
33. Verges, J. (1969) *Spectrochem. Act.* **24B**: 177.

Contents

9. Fourier Spectroscopy at the Air Force Cambridge Research Laboratories

A. T. Stair, Jr.
Optical Physics Laboratory
Air Force Cambridge Research Laboratories

Abstract

Recent and previously unpublished results of measurements using Fourier spectroscopy by various scientists of the Optical Physics Laboratory are presented. A major portion of the data comprises results of the airborne program on atmospheric emission and transmission of the atmosphere obtained at altitudes ranging from 8 to 12 km. The spectral region is from 1 to 15 micrometers. Spectra obtained of the total solar eclipse of March 7, 1970, are included.

9-1 INTRODUCTION

A primary mission of the Air Force Cambridge Research Laboratories is environmental research. This mission is directly translated by the Optical Physics Laboratory into measurements of the infrared properties of the atmosphere and laboratory studies of atmospheric species. A natural requirement, therefore, is also the continuous pursuit of better spectroscopic techniques, such as Fourier spectroscopy. Generally, the problems of interest do not require high resolution ($\Delta\sigma \lesssim 1\ cm^{-1}$) but rather are *energy limited* in the sense of time available and source brightness. Also, wavelength accuracy, pointing (field-of-view), and maximum signal-to-noise are very important. The atmospheric studies can be broken into about four categories: (1) thermal (equilibrium) properties studied both in transmission and emission; (2) nonthermal properties at high altitudes, such as infrared chemiluminescence; (3) auroral emissions; and (4) artificial perturbations such as high-altitude nuclear detonations.

The prime measurement platform is a four-engine jet aircraft, KC 135, 3120, with an altitude capability of about 12 km (40,000 ft) and a maximum mission time of about 10 hours. This aircraft is shown in Figure 9-1. The extensive modification of this airplane for optical measurements can be seen in the large number of special windows installed at various viewing angles. A major effort has been the development of techniques for using Fourier spectroscopy in such a hostile environment of vibration, temperature variations, and electrical noise.

9-2 SPECTRA OF HOT AIR FROM A HIGH ALTITUDE NUCLEAR DETONATION

The first successful measurements using interferometer spectrometers onboard the aircraft were made during the high-altitude nuclear test series conducted by the United States in 1962. H. P. Gauvin, Chief of the Radiation Effects Branch of the Optical Physics Laboratory, has the primary responsibility for these field efforts using the KC 135. At first it might seem that, of all things, measurements of a nuclear detonation should certainly not be *energy limited* and therefore not need the advantages of Fourier spectroscopy. Nevertheless, it happened in

Figure 9-1. Photograph of the AFCRL Optical Aircraft, KC-135

the reverse: A large f/4 Ebert-Fastie spectrometer with a 50-cm diameter mirror and 20×30-cm grating, which barely went through the cargo door in two pieces, obtained no data at all. In contrast, small interferometers with 1-cm diameter mirrors (the original Block Engineering cubes designed by L. Mertz) got excellent emission spectra, as shown in Figures 9-2 and 9-3. A significant advantage in this case was the larger fields-of-view of the interferometers since the relative positions of the aircraft and the target were impossible to predict a priori. In addition, however, for transient events the importance of *all* of the parameters of *energy limited* measurements are brought to the forefront, particularly the *time* available and the *resolution* required.

Figures 9-2 and 9-3 are spectra of a high-altitude fireball taken a few seconds after detonation. Each spectrum was obtained in one second. The instruments were vibrationally isolated by hanging them in rubber trampoline mounts. The primary features of both spectra are due to the three heteronuclear diatomic molecules expected in high temperature air: CO, NO, and NO^+. These data represent the first observation of infrared emission of NO^+; with the assumption of chemical and kinetic equilibrium, rough numbers have been obtained for the integrated band strengths of both the fundamental and first overtone of NO^+.* The resolution is about 35 cm^{-1}

* A. T. Stair and H. P. Gauvin, *Aurora and Airglow*, B. M. McCormac, Ed., Reinhold, New York, 1967, pp. 365–370.

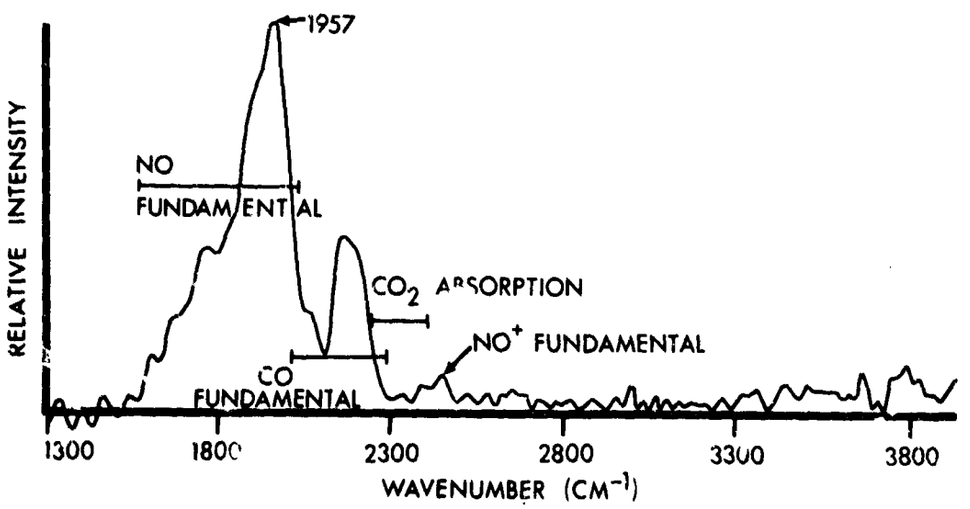


Figure 9-2. Spectrum of Hot Air (2.5 μ to 8 μ Region) From a High Altitude Nuclear Fireball. These data were obtained from 12-km altitude with an interferometer spectrometer using a bolometer detector

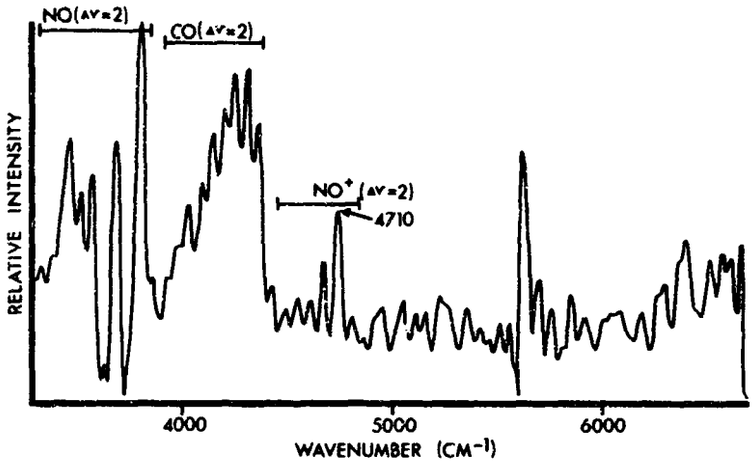


Figure 9-3. Spectrum of Hot Air (1.5 μ to 3 μ Region) From a High Altitude Nuclear Fireball. The data were obtained from 12-km altitude with an interferometer spectrometer using a PbS detector

(unapodized) which is not sufficient to resolve rotational lines, but it gives a very good picture of the high vibrational temperatures and the correct rotational and vibrational temperatures (not necessarily the same) can be determined by comparison with synthetic spectra. For example, the structure of the CO overtone of Figure 9-3 can be matched with a theoretical spectrum calculated at 3500° K ($T_{vibration} = T_{rotation}$) as shown in Figure 9-4. Marshall Bruce and James Rogers of this laboratory have compared many spectra of this resolution; in some cases the molecules are clearly in nonequilibrium.

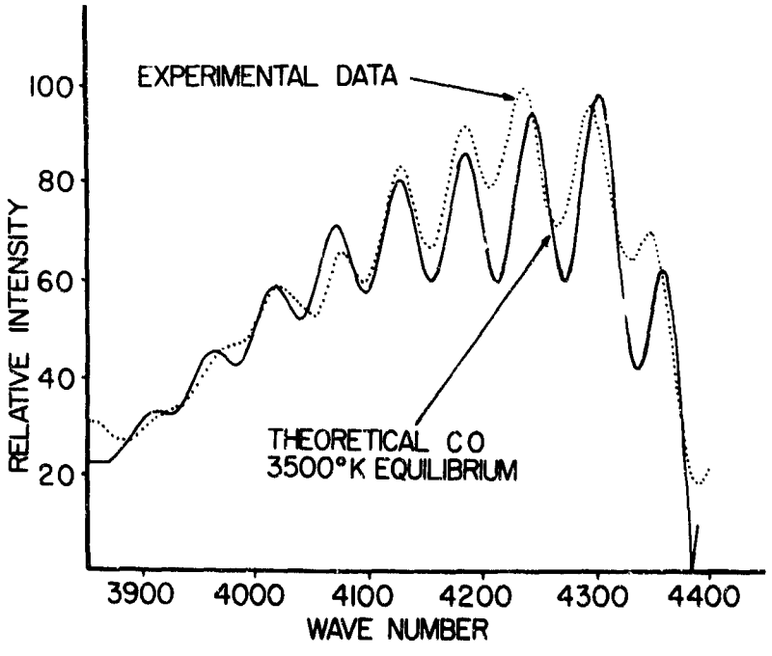


Figure 9-4. Comparison of the Overtone Spectrum of CO ($\Delta v = 2$) From a High Altitude Nuclear Detonation With a Theoretical Spectrum at 3500° K

The accuracy is typically better than 500° K. The overtone of nitric oxide can be seen to be severely absorbed by atmospheric CO_2 (2.7 μ) between the aircraft at 11 km and the hot air target at a much higher altitude.

9-3 ATMOSPHERIC TRANSMISSION MEASUREMENTS

For atmospheric transmission problems, the aircraft platform has certain advantages over ground based measurements. However, it has real limitations in many spectral regions (see Figures 9-2 and 9-3). Although excellent line by line computer programs now exist* for infrared transmission or emission calculations, the distributions of geography and altitude for the various species, particularly H_2O, are too little known to permit corrections. Consequently, a solar transmission measurement program using the KC-135 is also being pursued under the auspices of John Cahill of our laboratory. Interferometer spectrometers are used routinely for this work as well, although not truly required on an *energy limited* basis. However, the inherent advantages are fully utilized for pointing, accuracy, and better signal-to-noise, as well as the obtaining of data at different altitudes in a short time under comparable atmospheric conditions. During the past two years, transmission data have been obtained over a wide geographic area ranging from the North Pole to 55° S near South America and in the Pacific near Johnston Island and Hawaii.

* R. A. McClatchey, Optical Physics Laboratory, AFCRL.

The interferometer spectrometer is comprised of two back to back Block Engineering I4 cubes salvaged from old nuclear test instrumentation. The moving mirrors are physically connected and a HeNe (0.6328 μ) laser is pumped into the reference cube with the output used to command digitization. Also, white light is used with the reference cube in order to facilitate coherent adding of interferograms. One unit operates in the 1 to 3.3 μ region using an InAs (77° K) detector, and the other operates from 3 to 8 μ using Au(Ge) at 77° K. On the aircraft, both instruments are manually trainable covering elevation angles from +10° to +80°; they can be tracked ±12° fore and aft. The output signals are recorded analog on an Ampex CP 100 tape recorder. The aircraft window materials are GE 105 fused silica for the short wavelength instrument and CaF_2 for the other. The maximum optical displacement about the zero retardation point is approximately ±500 μ. This gives about 10 cm^{-1} resolution unapodized. The instruments are scanned rapidly, at a rate of 2 interferograms per second. This fast scanning coupled with Aeroflex wire rope shock mounts is sufficient to negate aircraft vibration problems. Both instruments were designed and fabricated by J. D. Rex of AFCRL.

Figure 9-5 is a representative sample of some of the reduced transmission data which have been corrected for the instrument response function and solar spectral variations. These results are for nearly zenith sun using the Au(Ge) detector. Approximately 50 interferograms have been coherently added. Tri-

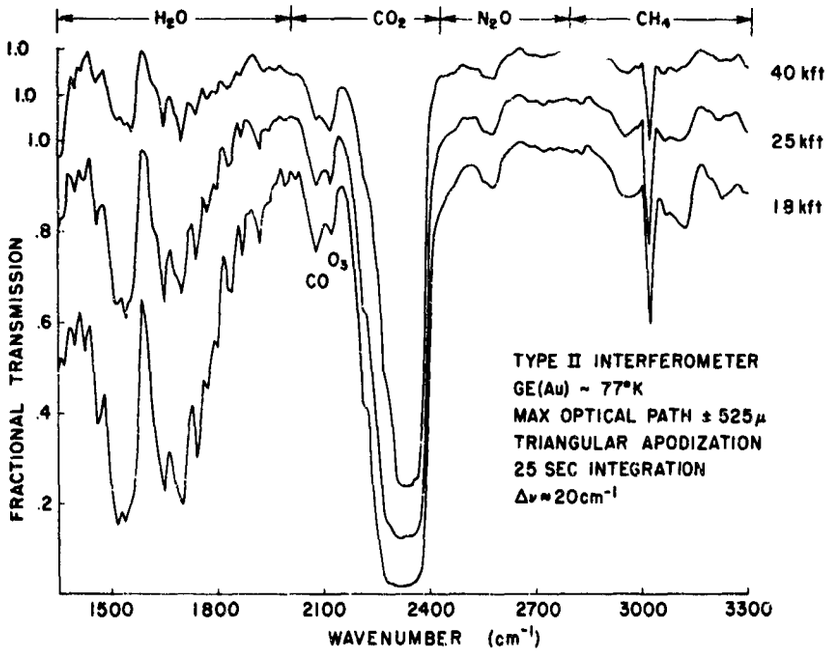


Figure 9-5. Fractional Transmission for Almost Vertical Path Through the Atmosphere From Three Altitudes: 18 000 ft; 25 000 ft; and 40 000 ft

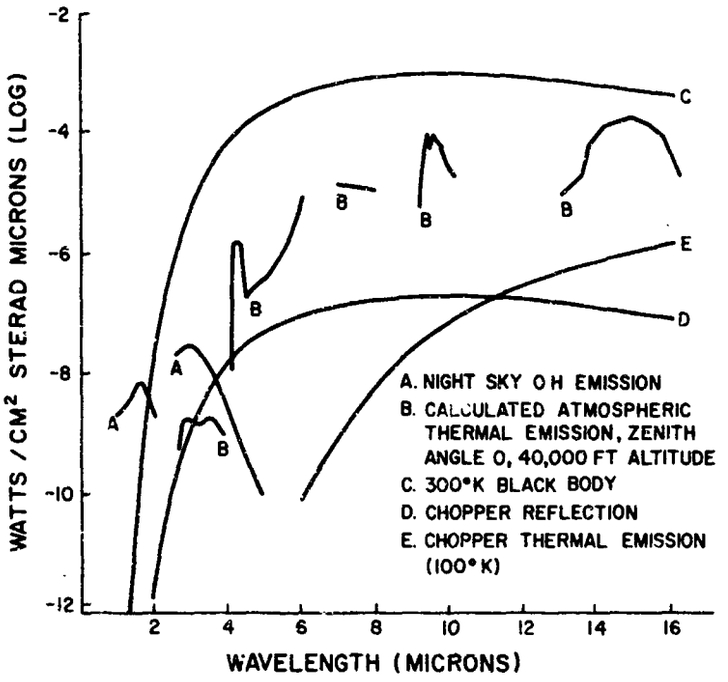


Figure 9-6. Zenith Atmospheric Emission Spectrum (Curves A and B) at 12-km Altitude Compared to Blackbody Emissions and the Chopper-Honeycomb Reflection Spectrum

angular apodization has been used yielding a resolution element of $\sim 20\ cm^{-1}$. The signal-to-noise ratio in the spectrum decreases toward smaller wavenumbers. At $3100\ cm^{-1}$, the s/n = 340; at $1600\ cm^{-1}$, the s/n = 100. The various atmospheric absorbers in the 3 to 8 μ region can be seen: H_2O, CO_2, CO, O_3, N_2O, and CH_4.

9-4 ATMOSPHERIC EMISSION MEASUREMENTS

The problem of obtaining infrared atmospheric emission spectra can be understood best by observation of Figure 9-6. Curves A and B are plotted as estimates of the spectral radiance of the atmosphere from 12 km, when viewed vertically. Curve C is the radiance of a 300° K blackbody which, if multiplied by some figure, say 5 percent, represents the level of radiation emitted by an aircraft window or the instrument optics when operated at ambient conditions. One solution, of course, would be to cool the entire optical system, including the aircraft window. An alternate approach which we have developed to get around this problem is to locate a liquid-nitrogen cooled chopper outside the aircraft window which is painted black to obtain about 1 percent reflectivity. The chopper is illustrated schematically in Figure 9-7. Three interferometers are mounted around the motor and operated simultaneously. With the blade cooled sufficiently, its emission becomes negligible relative to the atmospheric radiance levels. Curve E of Figure 9-6 illustrates this for the blade cooled to 100° K. The chopper is often operated as low as 70 °K. The limiting factor then becomes the radiation reflected off the blade from the hot optics inside the aircraft. To reduce reflected radiation even more, a field-limiting honeycomb, also cooled, is used to eliminate diffuse scatter from wide angles. Thus, another two-order decrease in reflected radiance levels is gained. This results in a reference chopper which has the apparent spectral radiance due to reflection of a 300° K black body reduced by a factor of $\sim 10^{-4}$, depicted as curve D in Figure 9-6. Thus, using phase sensitive electronics and liquid-nitrogen cooled detectors, InAs, InSb and He:Cd:Te, atmospheric spectral emission measurements at a zenith angle of 0° are possible in a few minutes to radiance levels of about $10^{-8}\ W\ cm^{-2}sr^{-1}\mu^{-1}$ at short wavelengths and $10^{-6}\ W\ cm^{-2}sr^{-1}\mu^{-1}$ at long wavelengths.

The foregoing statement, of course, is meaningful only in the context of many parameters: detector D*, time of measurement, resolution, throughput (AΩ) and a fudge factor, η, which includes such things as transmission and modulation efficiency (of both the interferometer and chopper system). The noise equivalent spectral radiance (NESR) for line sources using an interferometer in this mode can be shown to be:

$$NESR \simeq \frac{NEP}{\sqrt{t\phi\,\Delta\sigma\,\eta}},$$

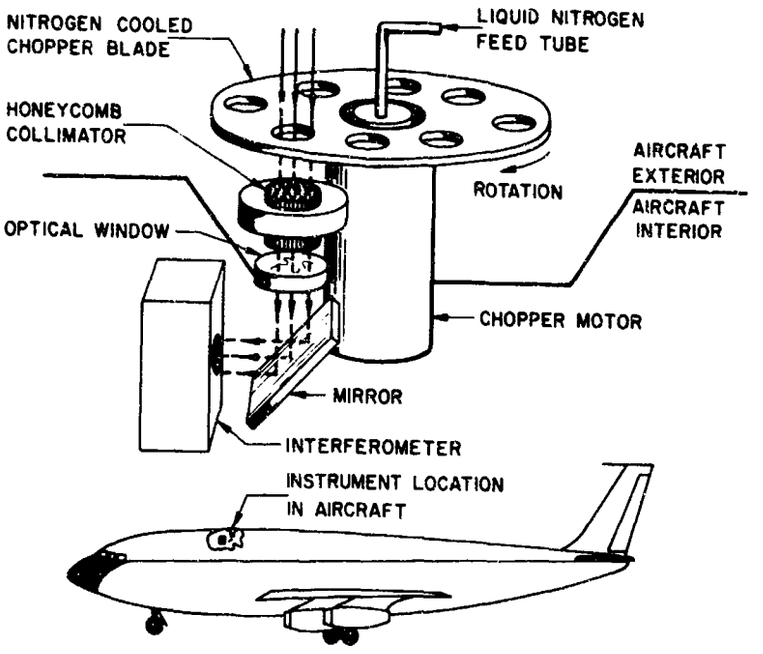


Figure 9-7. Schematic of the Exterior Liquid-Nitrogen cooled Chopper and Field-limiting Honeycomb. Not shown are two key items: (1) an aerodynamic screen located in front of the chopper system for raising the jet airstream over the assembly in order to leave a relatively quiet region; and (2) a cover plate with three holes for the three interferometers. This plate encloses the chopper wheel for safety and helps keep H_2O and CO_2 from the windows (frost) because of the continuous flushing action of the cold nitrogen gas

where

NEP is the noise equivalent power of the detectors, t is time, ϕ is throughput ($A\Omega$), $\Delta\sigma$ spectral resolution and η the efficiency factor.

As a typical example, consider an interferometer with an InSb detector.

$$\text{NEP} = 1\times10^{-12}\ \text{W}$$
$$t = 5\ \text{min} = 300\ \text{sec.}$$
$$\phi = 1.4\times10^{-2}\ \text{cm}^2\ \text{sr.}$$
$$\Delta\sigma = 5\ \text{cm}^{-1}$$
$$\eta = 0.025$$
$$\text{NESR}_{(5\mu)} = 3.2\times10^{-11}\ \text{W}/(\text{cm}^2\ \text{sr}\ \text{cm}^{-1})$$

The efficiency η is the product of several factors: the interferometer transmission and modulation efficiency, typically 0.1; the chopping efficiency, 0.35, and the honeycomb and window transmissions, 0.72.

Aside from the cryogenic chopper, the major problems have been the development of sufficiently accurate parallel ways for high resolution and the overcoming of extreme sensitivity to vibration. The Michelson interferometers were developed by Idealab, Inc., of Franklin, Mass., to maintain parallelism (one inch optics) to better than 0.8 sec of arc over 1 cm of drive. The vibration environment on the KC-135 was measured and found to be typically 0.2 g peak-to-peak, almost equally distributed over a bandwidth from a few to thousands of cycles per second. Rapid scanning is the best approach to the problem of vibration, providing large inertial forces and tighter servo operation. However, due to the thermal problem the chopper becomes the carrier frequency; it is rotated at the maximum subsonic speed (~1600 hertz with 8 holes). Consequently, the signal frequencies cannot exceed about 300 hertz. With these restrictions, the interferometers typically scan at 0.1 mm per second.

These instruments were significantly *hardened* by Idealab, and at this scan rate they operate adequately under about 0.1 g of vibration. All three interferometers are isolated to this level on a large plate which is the bottom of a sealed box, with temperature control to about 1° K, and with reduction of H_2O and CO_2 in the internal optical path by flushing with nitrogen. This chopper and interferometer system in entirety was designed and developed primarily by the author and E. Ray Huppi of AFCRL, with the assistance of Allan Steed and Ralph Haycock of Utah State University. The help of John LaSpina and Dr. Thomas Walter of this laboratory has been very instrumental both in obtaining and in reducing the data gathered with this system.

A block diagram with schematic representations of the various signal channels is shown in Figure 9-8. In

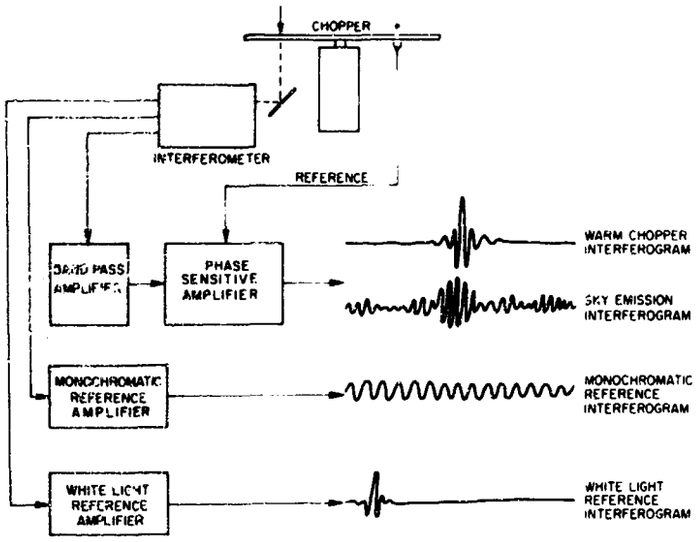


Figure 9-8. Schematic of the Electrical Signals Recorded. Phase sensitive amplification is used to discriminate against the large d-c interferogram due to thermal emission of the optics and internal scattered radiation

addition to the data channel, a monochromatic reference is necessary for proper sampling. For averaging the interferograms, a white light channel (strong central maximum) is used to identify the same monochromatic reference fringe on separate scans. When the chopper is at the ambient temperature its blackbody emission spectrum (continuum) produces an interferogram depicted by the upper curve of Figure 9-8. As the blade is cooled the signal decreases until a null is reached when the atmospheric radiant energy falling on the detector is exactly equal to that from the chopper. This is an excellent calibration of the atmospheric radiance since the blade temperature is carefully monitored and the paint is known to be very black in these spectral regions. As the blade cools the atmospheric emissions (which are 180° out

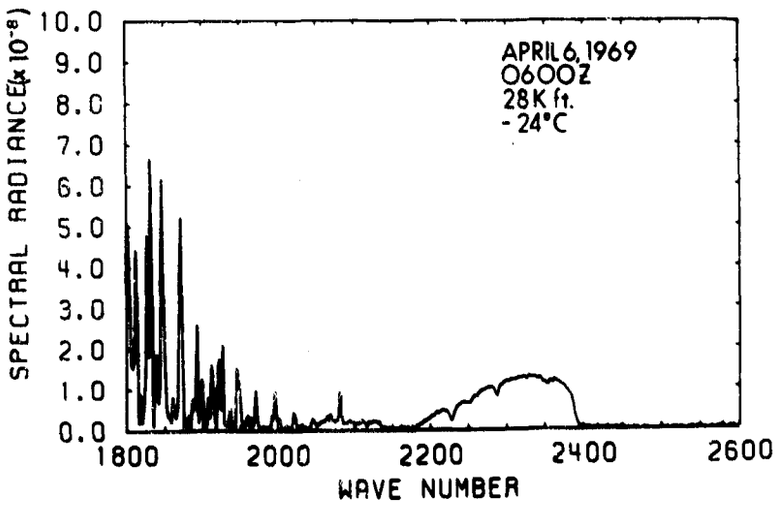


Figure 9-9. Atmospheric Emission Spectrum 4 μ to 5.5 μ, Zenith Sky, Dark, in Radiance Units of W/(cm² sr cm⁻¹) Taken From an Altitude of 8.8 km

of phase with the chopper radiation) produce interferograms with much more structure (2nd curve) which is typical for noncontinuum spectra.

Figure 9-9 is a nighttime atmospheric emission spectrum in the 4 to 5.5 μ region obtained from 8.8 km (28,000 ft). These measurements represent three minutes of observation time with an InSb (77° K) detector with a spectral resolution of about 2.5 cm^{-1}. The broad feature around 2300 cm^{-1} is the intense ν_3 band of CO_2. The maximum at 2210 cm^{-1} is the P branch of the ν_3 band of N_2O with band center minimum at about 2220 cm^{-1}. The features from 2130 to 2050 cm^{-1} are a combination of O_3 ($\nu_1+\nu_3$) around 2130 cm^{-1} and CO_2 ($\nu_1+\nu_2$) emissions, the sharp line at 2080 cm^{-1} being CO_2. The 3 ν_2 band of CO_2 is responsible for part of the structure around 1930 cm^{-1} with the other intense lines being due to the strong 6.3 μ band of H_2O (ν_2). In addition to the null calibration technique, external blackbody calibration sources are used to give the corrected spectral radiance in units of $W/cm^2sr\ cm^{-1}$. For the optically thick CO_2 band, this corresponds very closely with the measured outside air temperature. Figure 9-10 is a similar nighttime spectrum obtained from ~12 km (39,000 ft) where the over-all radiance level can be seen to be smaller; note that the local outside air temperature is also less, 234° K versus 249° K for Figure 9-9. In addition, it is clear that the H_2O emission is considerably less relative to CO_2 at higher altitudes in agreement with the expected difference in altitude distributions of these two species.

Figure 9-11 is an uncorrected spectrum obtained at night with a Hg:Cd:Te (77° K) detector and an interferometer using NaCl as the beamsplitter and compensator material. The dominant feature is the ν_3 fundamental band of O_3 centered around 1040 cm^{-1}. The detector response limited the short wavenumber spectrum to about 700 cm^{-1} which cuts out the major portion of the strong ν_2 fundamental of CO_2. Although the weaker ν_2 fundamental of O_3 ($\nu_2 \approx \frac{1}{20}\nu_3$) also occurs in the 700 cm^{-1} spectral region, the partially resolved band between 700 and 800 cm^{-1} can be identified with CO_2 in accordance with the relative abundance of CO_2 being 2 or 3 orders greater than O_3 up to very high altitudes. HNO_3 emission can be identified near 880 cm^{-1}. The other weak emission feature near 1300 cm^{-1} is where ν_4 of CH_4 and ν_1 of N_2O are known to overlap. These features have been identified better in other spectra not shown.

Data were also obtained during the recent total eclipse, March 7, 1970. Figure 9-12 shows spectra taken by J. Cahili and J. Rex of AFCRL with an interferometer spectrometer described under the section on transmission. The detector was InAs. The data are in a very preliminary state of reduction; for example, the instrument function and radiation balance corrections have not been applied. However, the gross coronal emission structure can be seen. The bottom curve represents the instrumental background as seen just after second contact for about 45 sec. During this time, the instrument was pointed at the dark eclipsed sky. The spectra seen here are negative; the optics are hotter than the source. The spectrum above 4900 and below 3200 cm^{-1} is noise, evidenced by a random distribution of the phase angle between the sin and cos transforms. The upper and lower graphs are on different absolute scales. At 45 sec after second contact the instrument was swung onto the corona. The field of view is ~5° circular diameter (to $\frac{1}{2}$ power) so that the bulk of the corona was observed.

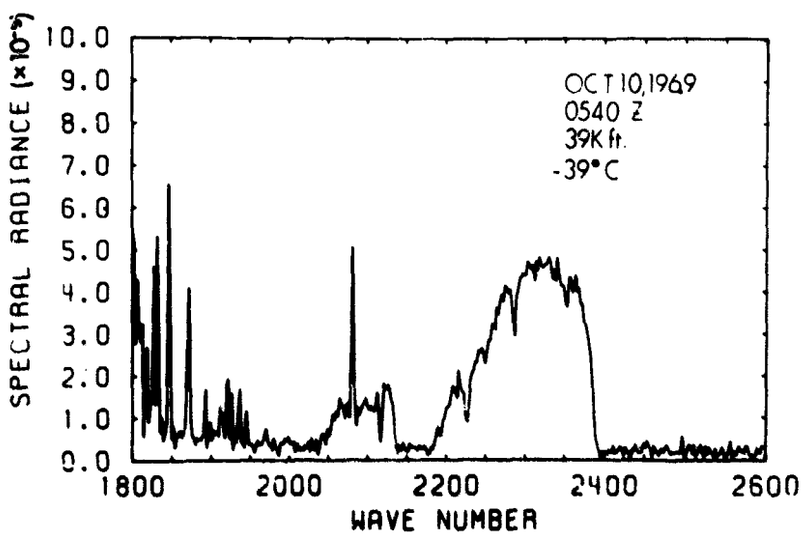


Figure 9-10. Atmospheric Emission Spectrum 4 μ to 5.5 μ, Zenith Sky, Dark, in Radiance Units of $W/(cm^2\ sr\ cm^{-1})$ Taken From an Altitude of 11.8 km

The negative region around 3700 cm^{-1} decreased in intensity, but still remained negative, indicating that the instrument was measuring a small influx of coronal energy in this spectral region. For wavenumbers greater than ~4100 cm^{-1}, a positive signal is observed. The upper curve resulted from the co-addition of 90 interferograms (2 scans/sec). Triangular apodization was used. Twenty interferograms

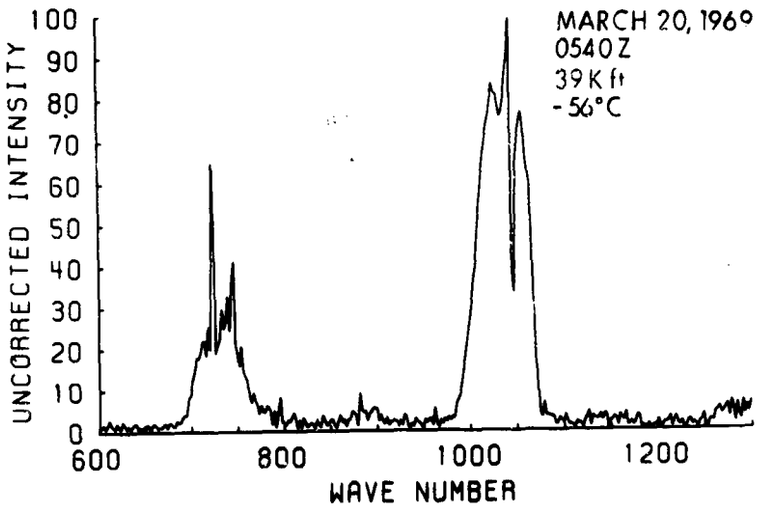


Figure 9-11. Atmospheric Emission Spectrum 7.5 μ to 14 μ, Zenith Sky, Dark, Taken From an Altitude of 11.8 km. The instrument response function has not been removed

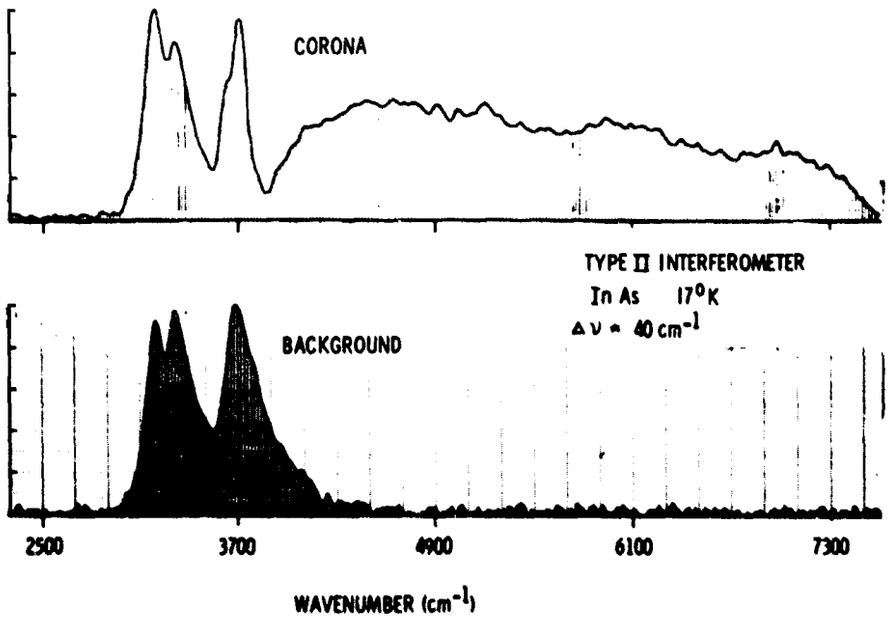
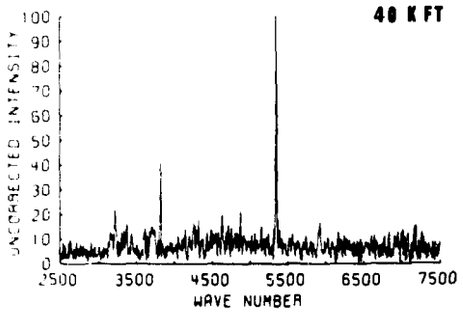
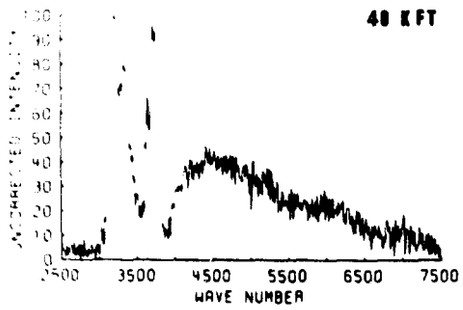


Figure 9-12. Emission Spectrum of the Solar Corona During the 7 March 1970 Eclipse (Upper Curve). The coronal spectrum is the result of co-adding 90 scans of $\frac{1}{4}$ sec each. The apparent background spectrum is due primarily to the thermal emission of the interferometer optics and aircraft window. Careful checking with phase information is required for correction of the coronal spectrum. This is being done (data shown were obtained only one week prior to reduction to this graphical form)

were used to obtain the background spectrum. The spectrum shorter than 2900 cm^{-1} (detector cutoff) is presented as a noise reference. The coronal emission appears to be a continuum with superimposed atmospheric absorption. The only apparent line structure is a relatively weak feature located at 6990 cm^{-1}. NASA also observed a weak feature at this frequency during the 1965 eclipse and tentatively identified it as Si X.



Solar Late Totality



Solar Mid Totality

Figure 9-13. Spectra of the Eclipsed Corona. The upper curve was obtained during a prominence that occurred a few seconds before third contact from one interferogram of ½ sec scan. The lower curve is similar to that shown in Figure 9-12, but results from co-adding only 50 spectral scans of the corona with no visibly apparent prominences.

Figure 9-13 is a spectrum obtained just before third contact when a solar prominence occurred and hydrogen Brackett and Paschen lines were observed. The lower curve is a repeated spectrum of coronal emission (50 co-added interferograms) for reference purposes. The upper curve is from a single spectral scan, ½ sec, taken during the flare. The two most prominent features in the upper spectrum are identifiable as hydrogen Paschen 4,3 at 5332 cm^{-1} and hydrogen Brackett 6,4 at 3808 cm^{-1}. Other spectral scans show more clearly the emission of Brackett 8,4 and 7,4 as well. The data are plotted with 20 cm^{-1} resolution, the instrument response has not been removed, and the spectra are normalized to the most intense feature.

9-5 LABORATORY MEASUREMENTS

Some of the laboratory applications of Fourier spectroscopy at AFCRL are represented by the papers published in this volume by Murphy and Sakai, Smith, and Hoffman. The paper by Murphy and Sakai is a new approach to gas kinetic studies which are more typically done by flowing afterglow techniques, such as depicted in Figure 9-14. The experiment is one of studying vibrational-vibrational energy exchange between vibrationally hot nitrogen ($N_2^{\ddagger}$) and CO. The requirement for interferometric spectroscopy can be seen from the fact that at typical operating pressures of 100 mTorr the total number of molecules in a 10 cm^3 volume is $\sim 3.5 \times 10^{16}$. Of these the number of excited CO molecules ($v \geq 1$) can be on the order of 10^{13} and smaller. Thus, the maximum radiance of the source (near the mixing region) can be approximately 10^{-7} W/(cm^2 sr cm^{-1}) and falling to 10^{-9} W/(cm^2 sr cm^{-1}) a meter downstream due to relaxation and radiative losses. These relaxation, quenching and radiative processes are the most interesting ones for study; therefore, spectra with good signal-to-noise are necessary. The experiment described here is also one of the most intense sources of this class.

Figure 9-15 is a spectrum obtained by J. P. Kennealy of this laboratory with the experimental arrangement just described. It is a spectrum of the first overtone of CO. Typically, these laboratory sources produce spectra of molecules with large vibrational excitation but rotational temperatures equal to the gas kinetic temperature since rotational relaxation occurs very fast (typically 3 to 10 collisions). Theoretically calculated spectra are used to determine the state of the gas, $T_{vib} = 5000°$ K and $T_{rot} = 350°$ K.

Other experiments such as $NO + O \rightarrow NO_2$ chemiluminescence are so feeble in the infrared that one not only needs Fourier spectroscopy but additional light-gathering techniques, such as infrared integrating spheres. This experimental arrangement is shown schematically in Figure 9-16. Another important point for all of these measurements is that the gas sources are almost always orders of magnitude weaker than the thermal emission of the environment. Consequently, one must, as a rule, modulate the reaction in some way and use phase sensitive a-c amplification to see the molecular emission. This is generally done, in our case, by modulating the microwave discharge which is the source of atoms or vibrationally excited species such as nitrogen.

9-6 SUMMARY

The primary thrust in the development of Fourier spectroscopy at AFCRL has been to apply it as an efficient tool at moderate resolution to a myriad of spectroscopic problems, primarily associated with

atmospheric physics. In addition to further work of the kind described here, there are a number of new Fourier instruments under development. These include the development of cryogenic interferometers for both aircraft and sounding rocket platforms to take advantage of the increased sensitivity of background limited (BLIP) infrared detectors. These are particularly important for upper atmospheric emissions, such as airglow and aurora, both for *energy limited* reasons and for getting above the optically thick thermal atmospheric emissions. Also, the wavelength range of airborne observations, primarily transmission (ambient temperature optics), will be extended to 1 mm.

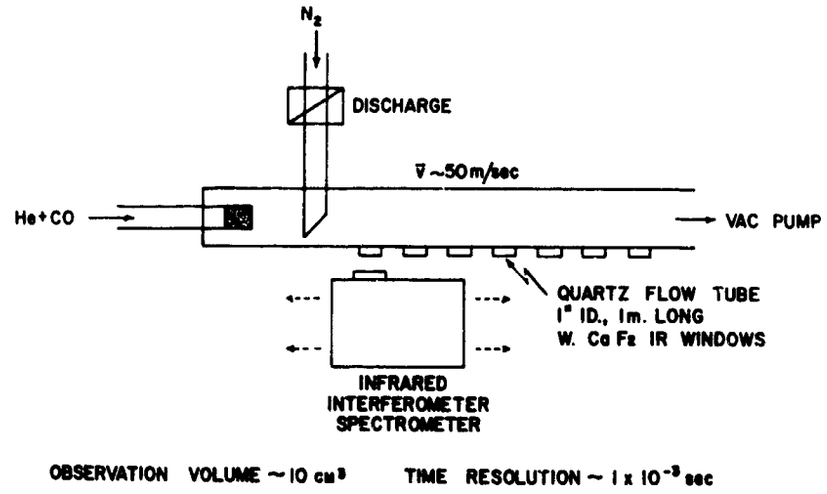


Figure 9-14. Schematic of a Laboratory Flowing Afterglow System for Gas Kinetic Energy Exchange Processes

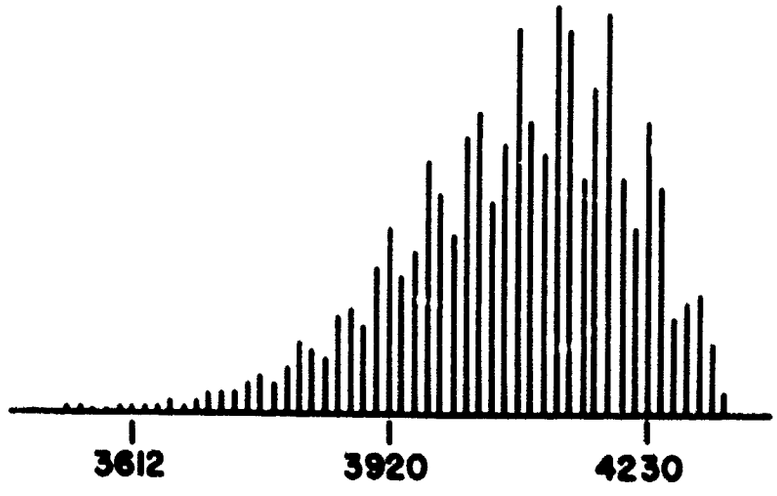


Figure 9-15. Spectrum of the First Overtone of CO Obtained by J. P. Kennealy From the Experimental Arrangement (Figure 9-14). Comparison with theoretical spectra shows the gas to be in a nonequilibrium distribution with a rotational temperature of 350° K but a vibrational temperature of 5000° K

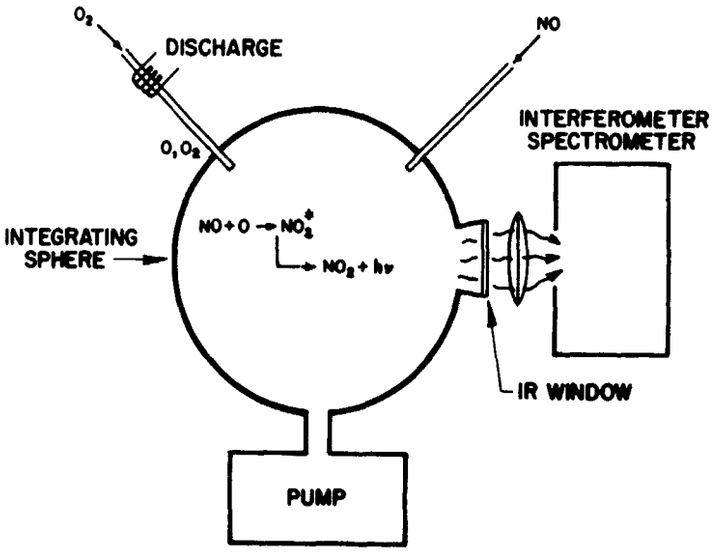


Figure 9-16. Schematic of a Flowing Afterglow System for Obtaining Spectra From a Source so Feeble That an Infrared Integrating Sphere is Required in Addition to Fourier Spectroscopy. These spheres are made by uniformly dimpling one or two liter Pyrex spheres, cleaning by ion bombardment, and coating with gold. Edwin Eckberg of this Laboratory is the superb craftsman of these spheres. The dimpling produces equal response to a point source independent of its location inside the sphere so that quantitative measurements are possible

Contents

10. The Origins and Logic of Multiplex, Fourier and Interferometric Methods in Spectrometry*

Peter Fellgett
Department of Applied Physical Sciences
The University of Reading
Reading, England

Abstract

The origins of interferometers, interferometric and Fourier spectrometry, and multiplexing applied either to spectra or to image resolution are traced. The theoretical and practical importance of the distinct attributes of multiplexing, Fourier methods, and interferometry applied to spectrometry are emphasized. Interferometry confers a luminosity advantage that can also be obtained by multislit methods (these have nothing to do directly with multiplexing, of course). Fourier methods confer the advantage of wide spectral range. Multiplexing confers a time advantage, equivalent to a signal-to-noise advantage. The greatest observational power results from applying all three separate attributes simultaneously in multiplex interferometric Fourier (MIF) spectrometry. A striking example shows that multiplexing is not the least of the three.

10–1 MONOCHROMATORS, MULTIPLEXING AND MICHELSON

There is evidence in the literature that much confusion still remains about the distinction between the multiplex, interferometric, and Fourier attributes in spectrometry, and the interrelation between these attributes. Clearly, if we are to understand the origins of these methods, we must first be clear about what is being discussed.

This is important for at least two reasons. The first is that an accurate understanding of abstract concepts is one of the essential components of the scientific method, and one on which the development of science depends in the long term. The second reason is that the distinctions are important even in relation to immediate practical applications, as we shall see.

*A taped version of this paper was played for the attendees at the conference.

Multiplexing is the process whereby a number of separate streams of information are sent along a single physical channel. It originated in telegraphy, has been extensively developed in modern trunk telephone circuits, and most recently has played an essential role in communications satellites. In multiplex spectrometry we pass information concerning a large number of separate spectral elements through a single detection channel. Multiplexing in optics is not, of course, confined to spectrometry. For example, Professor Jim Ring[1] and his colleagues have put it to good use in direct imaging applications.

In 1949 I was concerned with the possibility of gaining sensitivity by applying multiplexing in infrared spectrometry. The advantage of doing this is by now well known, and some authors have been kind enough to refer to it as "the Fellgett advantage". I think it is better physics to use descriptive terms rather than persons' names (for example, Kaiser and

Hertz) and would prefer to call it simply the multiplex advantage. It was first described in print by Dr. Marcel J. E. Golay,[2] quoting a private communication from myself. He puts it this way:

> The most objectionable feature of a monochromator in any spectrometric work in which radiation is at a premium is implied by the very name of the device: only one element of the spectrum is utilized at one time, while the bulk of the radiation is wasted on the jaws of the exit slit. When the source of radiation is small, a far more efficient use of the radiation can be made by employing the method now being developed by Fellgett for astronomical observations. This permits to obtain information about the complete spectrum in the time needed with the spectrometer to obtain one section of the spectrum corresponding to the slit width, when the entire source image is accepted at the monochromator entrance.
>
> This procedure bears a resemblance to Michelson's use of the visibility curve, although the circumstance that Michelson's visibility was essentially the auto-correlation function referred to some carrier in the center of the narrow spectrum observed, conferred a vectorial character to the visibility; this served to restrict the use of the visibility, obtained visually, and treated as a scalar quantity, to spectra which were symmetrical, or nearly symmetrical, with respect to the carrier. It must also be recalled that Michelson worked with visible light and placed emphasis on the accuracy yielded by his method. The advantage in signal-to-noise ratio, yielded by the interferometric method in the far infrared, and which does not exist, basically, in the visible region, has been first suggested by Fellgett, to the writer's knowledge.

In February, 1949, Dr. Golay had written to me:

> Michelson, under whom I took courses at the University of Chicago, developed his interferometer for the visible region, where there is no such advantage [in signal-to-noise ratio] basically, and he used to place emphasis on the accuracy advantages of his method. Also, he used to examine line spectra confined to a very narrow spectral region . . . .

Two points of special interest emerge from these quotations. The first is the reason why the multiplex advantage is not available in the visible region of the spectrum. There are in fact two reasons for this, of which one is extremely obvious. In the visible, image detectors such as the eye, photographic materials, and television camera tubes are available. There is accordingly no need to multiplex; we can look at all the spectral elements simultaneously with these imaging devices. The second reason depends more on the detailed physics. In the visible, the ambient radiation field is sufficiently weak at room temperature that the majority of photons falling on the detector are likely to come from the signal itself. Consequently, multiplexing a spectrometer does indeed give a signal-to-noise advantage, but there is a corresponding loss because of the increased noise in the detector; it is easy to show that the loss and gain are equal, so that they cancel each other out in typical circumstances.

The second conclusion from the quotations is that Michelson certainly did not formulate or exploit the multiplex principle. Neither indeed does he seem to have invented Fourier spectrometry in any explicit sense. Michelson's original publications suggest that he arrived at his understanding of the relationship between the spectrum of a light-source and the visibility of the interference fringes of the source by physical arguments based on beats between fringes produced by individual lines of the spectrum. The Fourier transform was apparently introduced by Lord Rayleigh in a letter to Michelson published in 1892.[3] It is interesting to note in passing that in this letter Rayleigh predicted what we should now call the rotational structure in molecular spectra; he, of course, did so on a classical-wave basis, since this was before the formulation of the quantum theory by Planck.

What Michelson did invent was the interferometer, and it is accordingly important to note exactly what he meant by this term. The phenomenon of interference had been known long before Michelson. His contribution was an instrument in which two interfering beams are well separated in space so that their relative path lengths can be conveniently and precisely varied. It is this that he calls an interferometer, a term which he appears to have coined. He is not, however, quite consistent about how restrictive the term should be. In *Light Waves and Their Uses*[4] he says:

> We may conveniently restrict the term *interferometer* to this arrangement, in which the division and the union of the pencils of light are affected by a transparent plane parallel plate.

However, earlier (*ibid.*, p. 33) he says:

> The utilization of the two portions of a lens, at opposite ends of a diameter, converts the telescope or microscope into an *interferometer.*

He here refers, of course, to what we now call the Michelson stellar interferometer. Having displayed some two dozen possible configurations involving either lenses or mirrors, or both, Michelson says (*ibid.*, p. 40):

> The form of interferometer which has proved most generally useful is that shown in Fig. 38.

And he then proceeds to describe the configuration which we traditionally associate with his name. Clearly it is this, and the stellar interferometer, which are properly called Michelson interferometers. The use of this designation for any two-beam interferometer is redundant, and fails to take account of the debt

we owe to Michelson for the very idea of an interferometer.

10–2 INTERFEROMETRY AND INVERSION

It was beyond the technical resources of Michelson's day, as Lord Rayleigh recognizes in the letter quoted, to measure the interferogram proper, and Michelson worked entirely with fringe-visibility curves. It is well-known that he conceived and built a most ingenious mechanical special-purpose computer for Fourier transformation. However, he is remarkably reticent about the precise use he made of this, and I have not found a positive statement that he successfully used it for the actual transformation of a visibility curve, although he attempted to do so (*ibid.*, p. 80), and successfully re-inverted a synthetic visibility curve (*ibid.*, p. 73).

The first reference I have found to the measurement of a genuine interferogram is to work by Rubens, Hollnagel and Wood in 1910–11.[5,6] Like Michelson, they guessed plausible spectral distributions and used semi-analytical methods to investigate the consistency of these distributions with the observed interferograms. Wood, writing in 1934,[7] confirms that Michelson used only transforms of guessed intensity distributions, and says that the full forward numerical transformation would have been "a much more difficult problem". To my surprise, it appears that I may have been the first person to do this, in the work at Cambridge already referred to,[8,9] although I was concerned with multiplexing and was using Fourier and interferometric methods only as a means to that end.

Michelson's invention of the interferometer, and his application of it to spectrometry, imply that he was the originator of interferometric spectrometry. His original method, which Rayleigh made into a Fourier method, was, however, somewhat of a *tour de force* which Wood[7] concluded had been little used since the original work. The main development of interferometric spectrometry was the non-multiplex, non-Fourier method associated with the names of Fabry and Perot[10] and of Lummer and Gehrcke.[11] It is unnecessary to describe these methods here, since they are treated in every textbook of physical optics.

The advantage of these interferometric methods was their large *luminosité* (product of area and solid angle of source accepted), but they had the disadvantage of very restricted free spectral range as compared, for example, to a so-called diffraction grating (a device which, strictly speaking, also works by interference). It was Professor Jacquinot[12] who first pointed out the advantages which could accrue by combining the luminosity advantage with the wide spectral range of Fourier methods.

By contrast, the stars I was attempting to measure in 1949 were so faint in relation to the available size of telescope and sensitivity of infrared detector that my concern was not with wide spectral range, but with getting a spectrum at all. Since the star-image is effectively a point source, it was not possible to exploit a *luminosité* advantage. The multiplex advantage was accordingly the only possibility open.

10–3 RECENT HISTORY AND REMAINING MISUNDERSTANDINGS

The rest of the story is sufficiently recent and well-known that it may be quickly told. By 1950 I had demonstrated a multiplex spectrometer, which happened also to be both Fourier and interferometric, and validated the multiplex advantage. The first person to make actual use of the multiplex method in applications appears to have been Dr. Larry Mertz in 1951, using a polarization method in the far infrared. Professor John Strong[13] also used the method in the far infrared, but apparently less for its signal-to-noise advantage than as a good method of discriminating against parasitic short-wave radiation. Madame Connes obtained a high-grade multiplex spectrum of the night sky in the near infrared in 1959. Since then, numerous developments have taken place, many of which were reviewed at the Bellevue Conference in 1966.[14]

Some investigators have unfortunately confused the multiplex, Fourier, and interferometric attributes, and have, consequently, not achieved the properties they were aiming at. The distinct and separate nature of these attributes has been illustrated by a diagram included in the *Proceedings of the Bellevue Conference* showing explicitly how every combination of the three attributes may occur. For example, the classic Fabry-Perot is interferometric but not multiplex or Fourier. The Mock interferometer is multiplex and Fourier, but not interferometric in the usual sense. Feeding a two-beam interferometer into a wave analyzer produces an instrument which is interferometric and Fourier, but not multiplex. This method has actually been used by experimenters who thought they were gaining a multiplex advantage.

It has to be emphasized that the multiplex property is a function of the entire system; neither the Fourier nor the interferometric attributes can guarantee it, nor are they necessary to it. Multiplexing is like virginity, if it is to be present at the end then it must be preserved throughout.

Multislit methods, aimed at increasing luminosity, have also been confused with multiplexing, despite Golay's clear statement already quoted[2] that his multislit spectrometer remains a monochromator. Dr. Martin Harwit[15] has recently demonstrated that binary coded slit-arrays can be devised which enable the multiplex and multislit attributes to be realized simultaneously. He rightly points out that these two attributes were implicitly combined by Fourier coding in the Mertz "mock interferometer,"[16] but this is the first time the possibility of this combination has (so far as I know) been made explicit and therefore brought fully into the design process. It is not, however, correct to refer to the combination as 'double multiplexing'; this phrase could logically refer to combined multiplexing of spectral and spatial information.

10–4 CONCLUSION

The most powerful combination of the multiplex, Fourier, and interferometric attributes to date has been in the justly famous work of the Connes[17] on the infrared spectra of planets. Each of the three attributes has been separately necessary to the feasibility of this work. The multiplex advantage is not the least among these. In recent laboratory work, the Connes have multiplexed one million spectral elements. This means that observations can be secured in a year which without the multiplex advantage, in an otherwise comparable spectrometer exploiting the Fourier and interferometric advantages, would take longer than man has been on this Earth.

References

1. Edgar. R. F. (1968) Some design considerations for infrared image scanning systems in *Infrared Physics, Vol. 8*, Pergamon Press, p. 183.

2. Golay, M. J. E. (1949) Multi-slit spectrometry, *J. Opt. Soc. Am.* **39** (No. 6): 437.

3. Rayleigh, Lord (1892) On the interference bands of approximately homogeneous light, *Phil. Mag.* **34** (No. 5): 407.

4. Michelson, A. A. (1902) *Light Waves and Their Uses*, University of Chicago Press. (Reprinted University of Chicago Press paperback (1961).

5. Rubens, H., and Hollnagel, H. (1910) *Phil. Mag.* **19** (No. 6): 761.

6. Rubens, H., and Wood, R. W. (1911) *Phil. Mag.* **21** (No. 6): 249.

7. Wood, R. W. (1934) *Physical Optics*, MacMillan, New York, p. 303.

8. Fellgett, P. B. (1951) Ph.D. Thesis, U. of Cambridge.

9. Fellgett, P. B. (1952) Proc. Ohio State symposium on molecular spectroscopy. (1958) *J. Phys. Radium* **19**: 187, 237.

10. Fabry, C., and Perot, A. (1897) *Ann. Chim. Phys.* **12** (No. 7): 459. (1899) *Ann. Chim. Phys.* **16** (No. 7): 115.

11. Lummer, O., and Gehrcke, E. (1903) *Ann. Physik* **10** (No. 4): 457.

12. Jacquinot, P. (1954) 17e Congrès du GAMS, Paris.

13. Strong, J. (1957) Frederick Ives Medal address, *J. Opt. Soc. Am.* **47**: 354.

14. Colloques Internationaux du Centre National de la Recherche Scientifique, No. 161, Orsay, 1966. (1967) Méthodes nouvelles de spectroscopie instrumentale, *Éd. du CNRS.*

15. Harwit, M., Phillips, P. G., Fine, T., and Sloane, N. J. A. (1969) Doubly multiplexed dispersive spectrometers, *CRSR 346*, Cornell University.

16. Mertz, L. (1965) *Transformations in Optics*, John Wiley, New York.

17. Connes, J., Connes, P., and Maillard, J. P. (1969) Atlas des spectre dans le proche infrarouge de Venus, Mars, Jupiter et Saturne, *Éd. du CNRS*, Paris.

Fourier Spectroscopy Past, Present, and Future

L. Mertz
Smithsonian Institute
Astrophysical Observatory
Cambridge, Massachusetts

This invited paper, presented during the conference, was published in the Feb. 1971 issue of Applied Optics.

Contributed Papers

Contents

11. Fourier Transformation of a 10^6 Samples Interferogram

H. Delouis
Centre Inter-Discipline Regional
De Calcul Electronique
C.N.R.S.
ORSAY, France

Abstract

The "decimation in time" method is applied to the discrete Fourier transform (DFT) of large interferograms when the number of samples is much greater than the available main storage size. A suitable technique for using the direct access disks in a 360/75 with 256 K of central storage has produced two different programs. The fast program, called "FFT 1024 K", gives the DFT of a $2 \cdot 10^6$ real samples function in 22 min; the second, called "TFD", gives the same result in 9 min in the case where the function is even or odd.

11-1 INTRODUCTION

Available fast Fourier transform (FFT) programs for complex, real, odd, or even functions can be applied only when the number of samples to be transformed fits into the computer central memory. For instance, with a 360/75 of 256 K words, it is practical to compute the Fourier transform of an interferogram with 128 K samples taken between $\delta = 0$ and $\delta = \delta$ max (J. Connes, Chapter 6). In order to transform 10^6 samples interferograms, programs using direct access auxiliary stores were developed; one is applicable to any $2 \cdot 10^6$ samples function, the other to a $2 \cdot 10^6$ even or odd function. In both instances, "decimation in time" was applied in order to break the Fourier transform operation into several elementary Fourier transforms, each fitting into the central memory, followed by suitable intercombinations.

11-2 DECIMATION IN TIME AND DFT OF A REAL FUNCTION—FFT 1024 K PROGRAM

The DFT of a real function I, sampled with $2N$ data is defined by

$$B_r = \sum_{k=0}^{2N-1} I_k W^{rk} \quad r = 0, 1, \ldots, 2N-1, \quad (11\text{-}1)$$

where $W = \exp(2\pi j/Nd)$ and $(j = \sqrt{-1})$. The $2N$ samples of the I function can be divided into two functions, Y_k and Z_k, each of which has only half as many points (N). The function Y_k is composed of the

even numbered points, and Z_k is composed of the odd numbered points:

$$Y_k = I_{2k}$$

$$Z_k = I_{2k+1} \quad k = 0, 1, 2, \ldots, N-1. \qquad (11\text{-}2)$$

Use the technique of decimation in time[1] to obtain B_r from the DFT of Y_k and Z_k.

If C_r and D_r are, respectively, the DFT of Y_k and Z_k, then

$$B_r = C_r + W^r D_r \quad 0 \le r < N.$$

$$B_{r+N} = C_r - W^r D_r \qquad (11\text{-}3)$$

This is illustrated in the signal flow graph of Figure 11-1 if $2N = 16$. In Figure 11-1, each B node is the

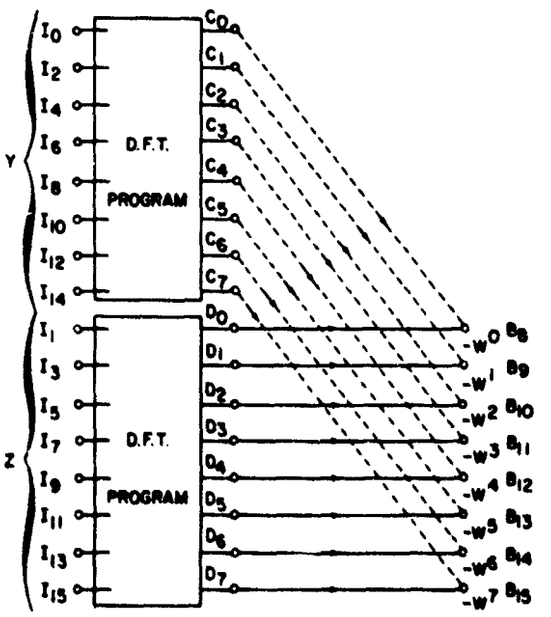


Figure 11-1. Discrete Fourier Transform Signal Flow Graph With $2N = 16$

sum of two terms: a D sample multiplied by a constant term $-W^n$ (solid arrow), and a C sample (not multiplied, represented by dashed arrow). For instance, $B_{13} = C_5 - W^5 D_5$.

As B_r is the DFT of a real function, $B_r = B^*_{2N-r}$, where B^* is the complex conjugate of B. Thus, only one half of the spectrum B is to be calculated.

For the same reason:

$$C_r = C^*_{N-r}$$

$$D_r = D^*_{N-r}. \qquad (11\text{-}4)$$

Replace the DFT program in Figure 11-1 by a FFT program called "FFT 64 K". This program was evolved at C.I.R.C.E. from the original Cooley-Tukey [2,3] algorithm, but it has been adapted to the DFT of a real function. As suggested by Bregham,[3] only one half of the complex spectrum must be calculated.

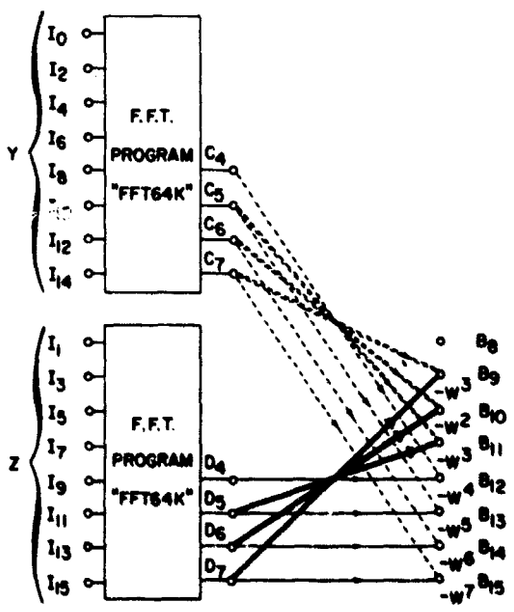


Figure 11-2. Fast Fourier Transform Program FFT 64 K

Figure 11-2 shows how Figure 11-1 must be modified in this case. The double dashed or solid arrow line indicates the transmission of the complex conjugate value of the variable written close to the original node. For example, $B_{11} = C_5^* - W^1 D_5^*$ and $B_{14} = C_6 - W^6 D_6$. Instead of having $B_9 = C_1 - W^1 D_1$ as in Figure 11-1, let $B_9 = C^*_{8-1} - W^1 D^*_{8-1}$ [see

ref. (3)], that is, $B_9 = C_7^* - W^1 D_7^*$. For the special case B_8, it is worth noticing that

$$B_8 = \sum_{k=0}^{15} I_k W^{8k} = \sum_{k=0}^{7} Y_k - \sum_{k=0}^{7} Z_k. \qquad (11\text{-}5)$$

Compare this with Eq. (11-2). B_8 is real.

Similarly, the computation technique shown in Figure 11-2 can be used to obtain C_r, $(r = 4, 5, 6, 7)$ from the N real samples Y_i. Thus, each time this technique is used, the number of samples to which the FFT program must be applied can be reduced by a factor of two.

Now, suppose that the main storage capacity of the computer is too small for computing the DFT of real functions sampled with more than four values, with the FFT program. The third signal flow graph shows (Figure 11-3) how it is still possible to compute the DFT of a $2N = 32$ sampled real function by using that FFT program. In the figure, a solid line brings a

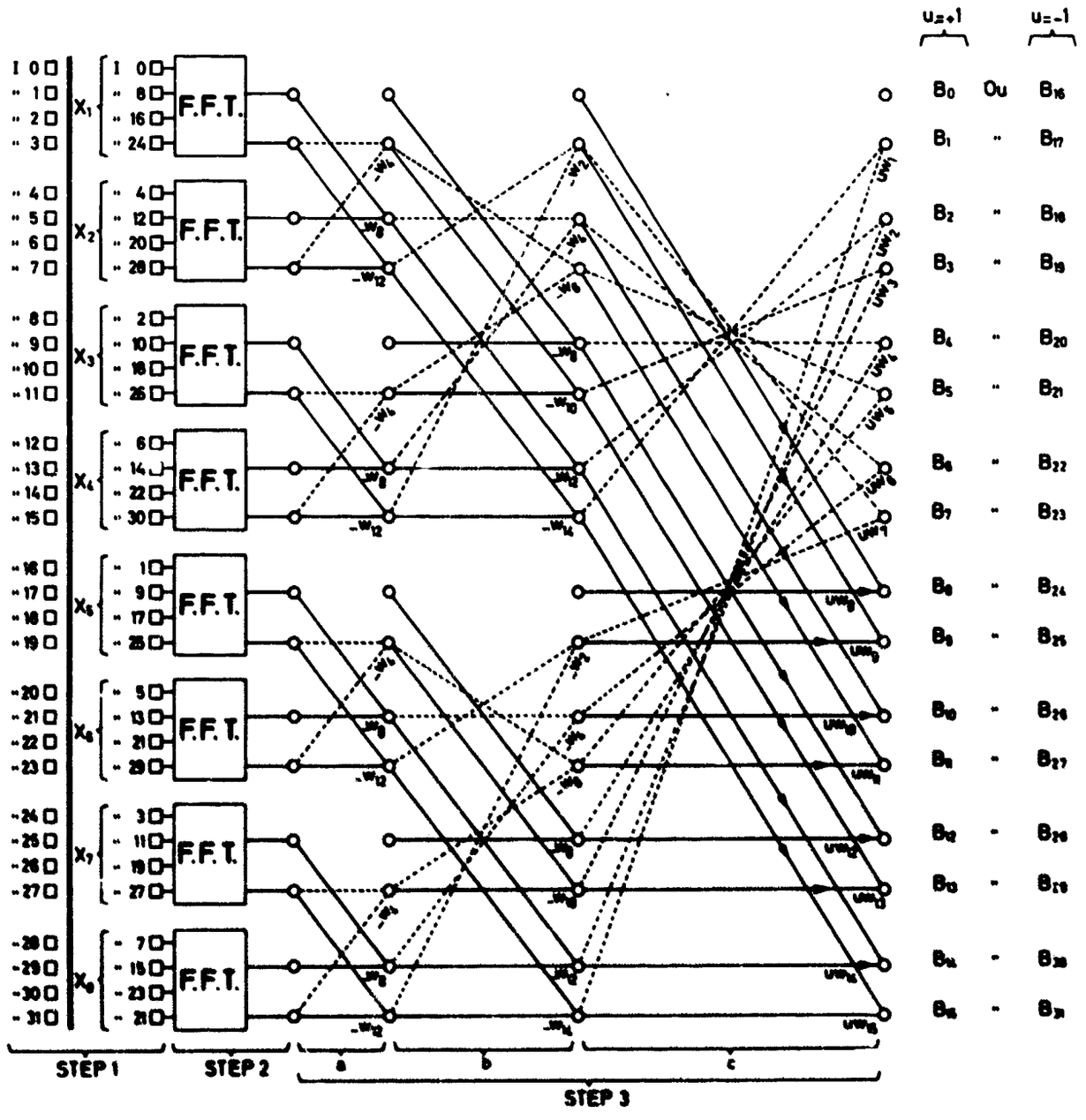


Figure 11-3. Computation of the Discrete Fourier Transform of a $2N = 32$ Sampled Real Function Using the FFT 64 K Program. $W^r = e^{2\pi j r/32}$

variable from the left to the right, while a dashed one brings the complex conjugate value of the variable represented by a node. Square nodes represent real values and round nodes represent complex values. $\sum(X_1+X_2)$ means the sum of all the elements of the two sequences X_1 and X_2, that is;

$$I_0+I_8+I_{16}+I_{24}+I_4+I_{12}+I_{20}+I_{28}.$$

In the first step, the original sequence of $2N$ real numbers I_k must be sorted into eight sequences $X_1, X_2, \ldots, X_8$ or $2N/8$ numbers. These eight sequences are stored within a direct access external storage device (for example magnetic disks). The FFT program can then be applied successively to each of these eight sequences within the main storage (step 2 on Figure 11-3). Thus, three successive steps are needed to obtain the final spectrum B. That last operation is clearly represented by Figure 11-3 (step 3), and the random nature of the transfers shows that it could not be made efficiently from sequential memories such as are provided by magnetic tapes.

11-3 APPLICATION OF THE FFT 1024 K PROGRAM

The purpose of this program is to compute the DFT of $2N=2^{21}=2048\,K$ (where $K=1024$) real points I_k on a 360/75 IBM computer having a main core storage of 256 K words. On this computer the best FFT program (called FFT 64 K) is available for a maximum of 128 K real samples, so by using the preceding method (Figure 11-3) the initial sequence I must be split into 16 subsequences X_i, each of which has 128 K points. Since $16=2^4$, four handling steps (instead of three as in Figure 11-3) are necessary to reconstruct the $2N$ values of the spectrum B from the X_i's FFT.

Except for the DFT computation of the 16 subsequences X_i which is made in the main storage, all the handlings are made by using eight data sets on two disk-packs of an IBM 2314 direct access disk storage unit. This device allows the search of randomly located data records the length of which have been limited to 410 words because of the main storage capacity. (A double buffer of 410 words must be reserved in the main storage for each of the eight data sets.) The access time to these records has been reduced by blocking them so that a read statement in the program at once brings 20 records (that is to say 8 K words) from the disk-pack to the main storage where the handlings can then be made. On one of the two disk-packs, 6780 records (or 1920 tracks) are to be reserved and on the other disk pack, 5120 records (or 1280 tracks). There are 4000 tracks on one disk pack.

11-4 DFT OF A REAL ODD FUNCTION WITH THE TFD PROGRAM

This program makes use of the method described by Brenner[4] and by J. Connes in Chapter 6 for computing the Fourier transform of a real odd function. Recall briefly the essentials.

To compute the Fourier transform B of a real odd Function I:

$$I_k=-I_{2N-k} \quad k=(0, 1, 2, \ldots, 2N-1). \quad (11\text{-}6)$$

First define a complex function G such as:

$$G_k=I_{2k+1}+iI_{2k}, \quad k=(0, 1, 2, \ldots, N-1). \quad (11\text{-}7)$$

Then the Hermitian function H should be

$$H_k=\tfrac{1}{2}(G_k-G^*_{N-k}) \quad k=(0, 1, \ldots, N-1). \quad (11\text{-}8)$$

The Fourier transform is a real function K_r. In order to compute K_r from the N complex samples H_k, define a new function F with only $N/2$ complex samples F_k:

$$F_k=H_k+H^*_{N/2-k}+i(H_k-H^*_{N/2-k})W^k. \quad (11\text{-}9)$$

Let E_r be the DFT of F, $E_r=\text{DFT}\,(F_k)$, and $r=(0, 1, \ldots, N/2-1)$. It is now possible to compute the real function K_r, since

$$E_r=K_{2r}+iK_{2r+1}. \quad (11\text{-}10)$$

K can be obtained by a simple sorting operation,

$$K_r, \quad r=(0, 1, \ldots, N-1). \quad (11\text{-}11)$$

One of the two halves of the spectrum B, which is purely imaginary, can be obtained from

$$\begin{cases} B_r = \dfrac{i}{2}\left(\dfrac{K_r + K_{N-r}}{\sin\dfrac{\pi r}{N}} - K_r + K_{N-r}\right) \\ B_{N+r} = \dfrac{i}{2}\left(\dfrac{K_r + K_{N-r}}{-\sin\dfrac{\pi r}{N}} - K_r + K_{N-r}\right) \end{cases} \cdot \qquad (11\text{-}12)$$

The TFD program has been adapted for computing an odd interferogram sampled from $\delta = 0$, with $N \leq 10^6$ on the 360/75 computer of C.I.R.C.E.; the main storage capacity is 256 K words. One disk storage unit was used in this case too.

The decimation in time technique is used for computing E_r from F_k. The set of E samples is broken into eight sub-sets of $N/16$ complex samples each; these are stored on one disk unit. The FFT 32 K program 1, which is available for a maximum of 64 K complex samples, is then applied to each subset in succession; the computation is performed within the central memory and gives the eight DFT's. Only three data sets on just one disk pack are needed. For $N = 10^6$ each data set contains only 10^5 words.

11-5 CONCLUSION

These programs could also be very useful with smaller computers having a disk storage unit for handling shorter interferograms. Trials of elapsed time and main storage total requirements have been carried out with the 360/75 of C.I.R.C.E. Table 11-1 gives the results when computing the DFT of symmetric real interferograms starting from zero path difference and sampled at N points. (With the FFT 1024 program, N input data points each equal to zero must be added.) These N points are read as input data from a magnetic tape, and the initial sorting step of the program only requires one reading of this tape. At the end of the program, the spectral samples are stored on a disk pack. Another program then computes four interpolated spectral samples between the Fourier transformed ones; thus, a smooth curve can be traced by a regular curve plotter.

Table 11-1. Computation Characteristics of FFT 1024 K and TFD Programs*

Case I $N_d = 2N =$ 1.K = 1024	2048 K	1024 K	512 K	256 K	128 K	64 K	32 K
Case II $N =$	1024 K	512 K	256 K	128 K	64 K	32 K	16 K
FFT 1024 K Elapsed time = Case I or II	22 min 8 sec	11 min	5 min 27 sec	2 min 46 sec	1 min 24 sec	43 sec	23 sec
TFD Elapsed time = Case II only	9 min 11 sec	4 min 35 sec	2 min 16 sec	1 min 08 sec	34 sec		
Main storage necessaries (K words)	180 K	100 K	60 K	40 K	30 K	25 K	23 K

*Case I is on N_d points sampled real function and Case II is on $N = N_d/2$ points sampled real odd function starting from zero path difference. Main storage necessaries = $\frac{8N}{33}$ + Disk buffers + program area (program area = 13,000 words). Magnetic disk area needed by: FFT 1024 K → $8N$ words (2300 tracks), TFD → $1.3N$ (1024 tracks if $N = 10^6$).

References

1. Cochram, W. T., Cooley, J. W., Favin, D. L. et al. (1967) *Proc. IEEE* **55**: 1664.

2. Cooley, S. W. and Tukey, J. W. (1965) *Mathematics of Computation* **19**: 297.

3. Bregham, E. O., and Morrow, R. E. (1967) *IEEE Spectrum* **63**.

4. Brenner, N. W. (1968) Private communication.

Contents

12. Fourier Spectroscopy From 10^6 Samples

Guy Guelachvili and Jean-Pierre Maillard
Laboratoire Aimé Cotton, C.N.R.S.
Orsay, France

Abstract

Previously demonstrated very high resolving powers have been extended to broad spectral ranges as well with a new interferometer. This has an increased sampling speed and all the recording operations are servo-controlled. The servo system uses phase modulation of interference fringes. The first results include rare earth emission spectra (thorium and holmium from 12500 to 3000 cm^{-1} with 0.02 cm^{-1} resolution) and absorption spectra of ICH_3 and N_2O from 3700 to 5600 cm^{-1} with 0.005 cm^{-1} resolution.

12–1 INTRODUCTION

The first high-resolution interferometer constructed at the Aimé Cotton Laboratory[1] was limited to a very small spectral range because it was well matched to the computation capabilities of a few years ago. But the advances made possible by the new computer generation have changed the situation. Thus, it is now practicable to obtain a Fourier spectrum from 10^6 samples in 9 min.

Consequently, it was necessary to design a new interferometer to exploit these recent possibilities. By this fact one achieves an important step in taking better advantage of the multiplex gain.

We will describe this new device, which features fast sampling speed and large flexibility in the choice of the step length and the internal modulation amplitude. The operations of displacement, positioning, and internal modulation are continually servo-controlled. The system is relatively insensitive to vibrations and intensity fluctuations of the reference line, which is the 3.50 μ superradiant line of Xe. Operation is thus made much safer. Main stress will be placed on the results, and only a summary description of the system will be given here; a more detailed one is found elsewhere.[1]

12–2 GENERAL DESCRIPTION OF THE INTERFEROMETER

A Michelson type interferometer equipped with two cat's eyes (Figure 12-1) was used. The beam has a diameter of 80 mm (about 3 in.). The path difference can reach 2 m, which is large enough for the majority of spectroscopic problems. The beam-splitter is made of two plane, parallel, equally thick plates flat to one fourth of a visible wavelength. Several pairs with suitable substrates and coatings are provided for each spectral range (visible, near infrared, middle infrared). A linear motor with 5 kgm of maximum power moves one of the cat's eyes which slides on oil bearings. Two piezoelectric ceramics supporting the secondary mirror of the cat's eyes permit very fast path-difference corrections.

12–3 PRINCIPLE OF THE SERVO-SYSTEM

An interferometer can be considered as an amplitude modulator. The intensity of the emerging signal depends on path difference. The disadvantages of this method for measuring path difference are well-known. Consequently, a phase modulation system was used.

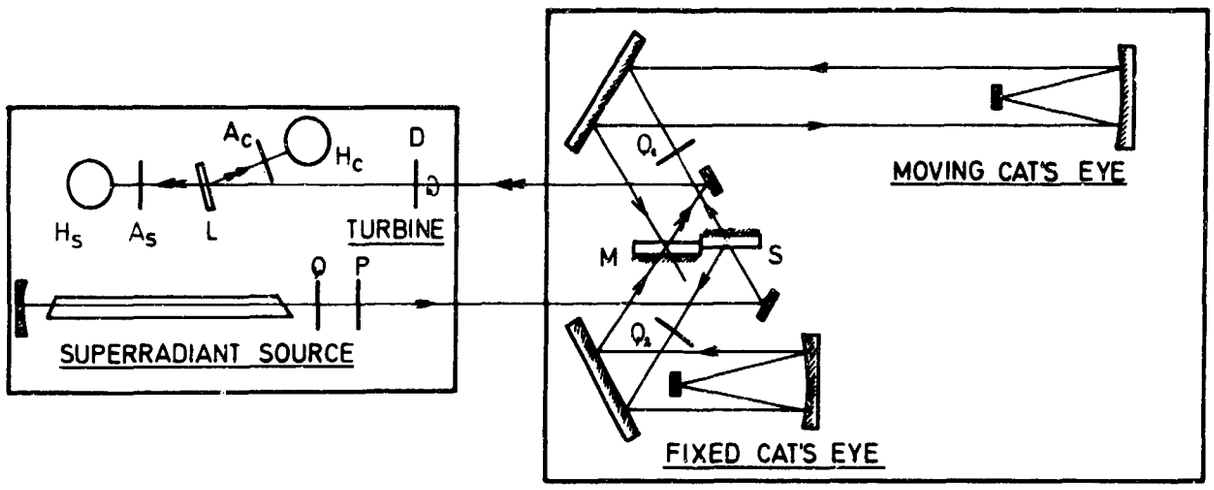


Figure 12-1. Optical Diagram of the Interferometer Showing Only the Reference Light Path, Which Utilizes the Center of the Beam Emerging From the Source Under Study

12-3.1 Phase Modulation System

The reference line λ_R, emerging from the superradiant tube (Figure 12-1), is linearly polarized by P before entering the interferometer. The axis of the polarizer is set at 45 deg to the axes of two quarter-wave plates (Q_1 and Q_2) located in each arm. They give two counter-rotating circularly polarized waves. They interfere and produce a linearly polarized vibration. Its azimuth α is a function of path difference Δ according to the formula

$$\alpha - \alpha_0 = \pi\left(\frac{\Delta - \Delta_0}{\lambda_R}\right). \qquad (12\text{-}1)$$

This light passes through a half-wave plate rotating at frequency N, and falls on an indium antimonide photovoltaic detector H_s in front of which is placed an analyzer A_s. Thus, one obtains a $4N$ frequency signal F_s and

$$F_s = I \sin 2\pi\left(4Nt + \frac{\Delta - \Delta_0}{\lambda_R}\right). \qquad (12\text{-}2)$$

F_s is the carrier wave. Its phase depends on path difference. Now, if one assumes that Δ varies linearly

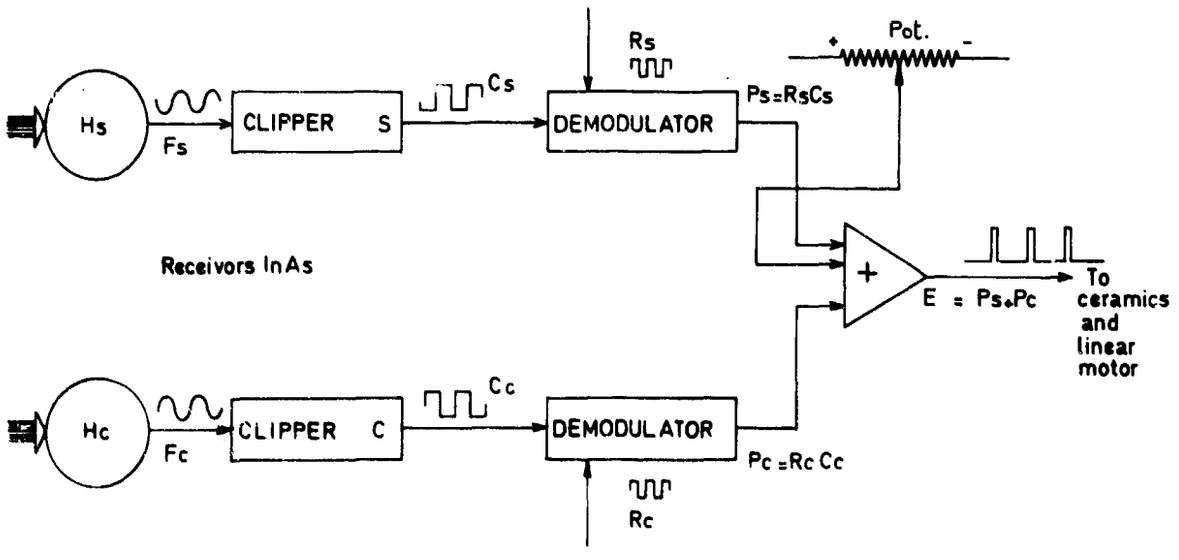


Figure 12-2. Block Diagram of System Producing Error Signal From F_s and F_c and the Reference Signals R_s and R_c. The potentiometer is used to add a direct voltage to the error signal, producing a variation in path difference equal to any fraction of the elementary step ($\lambda_R/100$). It is needed to adjust the first sample at zero path difference

as a function of time according to $\Delta-\Delta_0=Vt$, F_s becomes

$$F_s=I\sin 2\pi\left(4N+\frac{V}{\lambda_R}\right)t. \qquad (12\text{-}3)$$

Thus, frequency change is proportional to the speed of the carriage (single side-band modulation).

Detection of the F_s phase gives information on the path difference and the frequency itself yields a damping signal.

12–3.2 Error Signal Generation (Figure 12–2)

12–3.2.1 PRINCIPLE

After amplifying and clipping, F_s becomes a constant amplitude square wave C_s. In a synchronous demodulator C_s is multiplied by a reference signal R_s, the frequency of which is also $4N$, and the phase adjustable. The mean value of the product R_sC_s depends on phase difference between R_s and C_s. For a given phase of R_s, the variation of $\overline{R_sC_s}$ can represent the error signal. If it is plotted as a function of path difference, one obtains a triangular function with a period equal to λ_R (Figure 12-3). But there is a

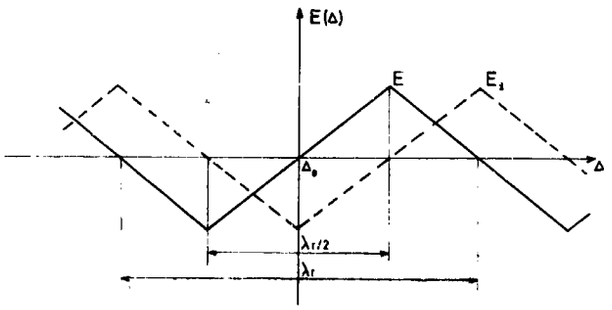


Figure 12-3. Error Signal Amplitude as a Function of the Path Difference Near Δ_0. The phase of reference signals R_s and R_c is stationary

troublesome residual ripple at $8N$ frequency when $\overline{R_sC_s}=0$. In order to eliminate this residual, the reference beam is split after the modulator (by L, see Figure 12-1) into two parts which give two signals in quadrature C_s and C_c. C_c is multiplied by R_c, a reference wave in quadrature with R_s. Finally, the sum $E=R_sC_s+R_cC_c$ is the error signal; it is made up of a train of rectangular pulses (Figure 12-4). The width of each pulse is proportional to the phase difference between C_s and R_s (or C_c and R_c). Inversely, by changing the phase of R_s and R_c it is possible to induce a variation of path difference

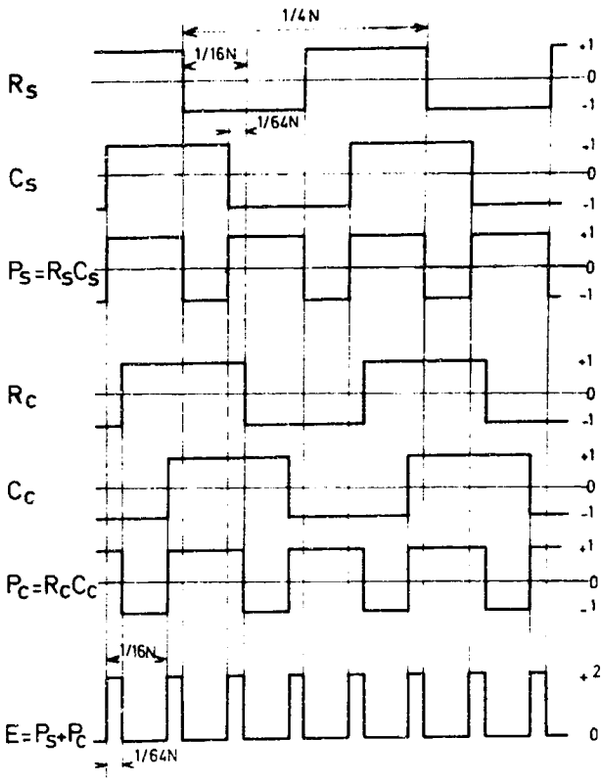


Figure 12-4. Waveforms Leading to the Error Signal E. In the illustrated case $\Delta-\Delta_0=\lambda_R/16$

because of the servo action. The $16N$ frequency of the pulses means that the minimum response time of the servo system is $\frac{1}{16N}$.

12–3.2.2 DESCRIPTION (Figure 12-5)

To produce R_s and R_c, a disk V_A with 400 radial opaque lines is attached to half-wave plate D. It rotates in front of a similar fixed disk V_B and chops two white light beams which generate a $400N$ frequency signal. This is frequency divided by a factor 100 in a reversible counter; the two outputs, of frequency $400N/100=4N$, are R_c and R_s. Thus, the phase of R_s and R_c can change by multiples of $2\pi/100$.

By suitable programming, one can obtain any variation of the path difference as a function of time. Therefore, to record an interferogram by the stepping technique, and using internal modulation at the same time, the law given by Figure 12-6 is used.

12–3.3 Recording System (Figure 12–7)

The samples are stored by an incremental magnetic tape recorder in a six character BCD form. A manual attenuator changes the gain of the signal in order to increase the dynamic range of the system. To correct the intensity fluctuations of the source, a part of the light under study is used to servo-control the gain of an amplifier before the demodulator.

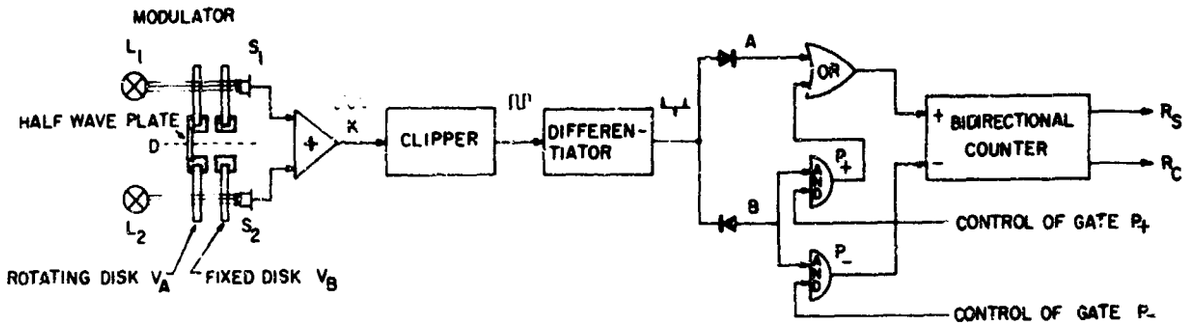


Figure 12-5. Production of the Reference Signals R_s and R_c

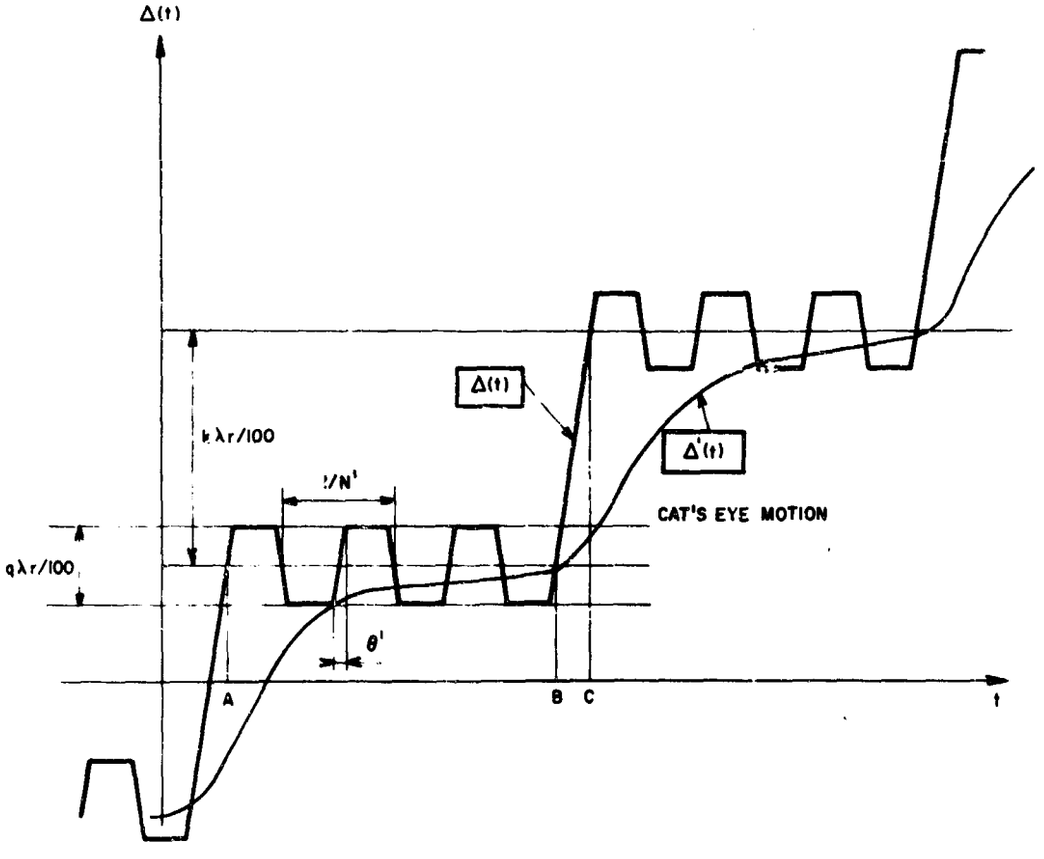


Figure 12-6. Path Difference Variation vs Time Needed to Record an Interferogram With Stepping Plus Internal Modulation. The approximate carriage displacement is given by $\Delta'(t)$. The piezoelectric ceramics supporting the small mirrors provide the difference $\Delta - \Delta'$

12-4 PERFORMANCES AND ADVANTAGES OF THE SYSTEM

12-4.1 Performances

The half-wave plate can rotate up to $N = 1000$ t/sec; that means that the minimum delay time of the servo system is 62 μ sec. In fact, the time limitation is due to the response time of the ceramics carrying the small mirrors. The actual response time is about 0.5 msec.

The piezoelectric ceramics give a 20 μ elongation for 2 kV voltage. This means they are able to provide a variation of $\pm 80\ \mu$ in the path difference. They also give a nearly square wave internal modulation; this is the optimum shape.

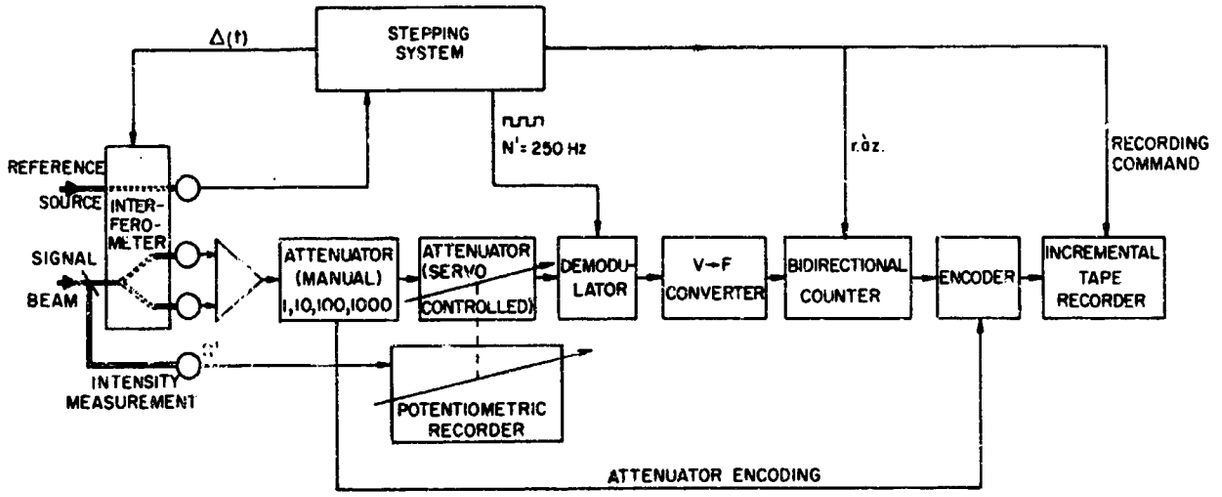


Figure 12-7. Complete Recording System Block Diagram

The maximum recording speed is presently limited by the digital recorder (Digi-Data Inc.) to about 50 samples/sec (with six BCD characters).

For an interferogram, the total time to actual measuring time ratio is about 0.8 in the worst case.

12-4.2 Advantages

The more evident advantage, compared to the previous systems, is the shortness of the elementary step. It is equal to $\frac{1}{100}\lambda_R$, that is, 350 Å for the reference line used.

The displacements and the amplitude of the internal modulation are integer multiples of the elementary step. Thus, there is a large flexibility in the choice of the interferogram sampling. For a given spectral range the suitable step can be selected practically according to the sampling theorem.

Moreover, the pulses which command the variation of path difference are independent of the reference fringes. The interference fringes are used for servoing but are not counted; the safety of operation is much increased.

The servo system has a large perfectly linear range: $\pm\frac{\lambda_R}{4}$ in path difference, that is, $\pm 0.88\,\mu$ in our case. For the previous systems using the measurement of fringe intensity,[1,3,4] the useful linear range was $\pm 0.1\,\mu$ only.

12-5 RESULTS

An attempt was made to obtain a great number of samples within a single interferogram, allowing the computation of a very wide spectral range. The same resolution ($d\sigma = 5{\cdot}10^{-3}$ cm^{-1}) as the resolution of the interferometer built by Pinard[1] was chosen, since it appears good enough for a multitude of spectroscopic problems considering the numerous causes of the spectral line broadening.

In principle, the apparatus is able to give spectra including more than $4{\cdot}10^6$ spectral elements. This number corresponds to the spectral range of the PbS receiver, and a recorded resolution equal to $5{\cdot}10^{-3}$ cm^{-1}.

Most of the difficulties involved are well expressed by the quality factor $Q = \frac{MS_m}{B_{eff}}$,[3] which is the ratio between the energy of the whole recorded spectrum and the energy of a spectral line which can be just distinguished from the noise. (M = number of spectral elements, S_m = mean value of energy per spectral element, B_{eff} = rms noise level.)

Interferograms from emission and absorption spectra have been recorded with PbS receivers.

12-5.1 Emission Spectra

Interferograms have been recorded from thorium and holmium sources. They are sealed quartz tubes containing the iodide of the rare earth. HF excitation (2450 MHz) is used. For the two elements, the studied spectral range covers all the sensitivity domain of the PbS receivers. The spectra are limited on the long wavelength side by the response of the detector at 3000 cm [illegible] the visible side by a filter opaque below 0.8 μ (... 0 cm^{-1}).

The length of the steps is equal to $0.455\,\mu$ (13×350 Å). This length corresponds to a free spectral range from 0 to 11,000 cm^{-1}. The peak-to-peak amplitude of the internal modulation

$$(0.70\,\mu = 20 \times 350\ \text{Å})$$

has been selected to have the maximum efficiency at 6150 cm^{-1}. In both cases (Th and Ho) the recording

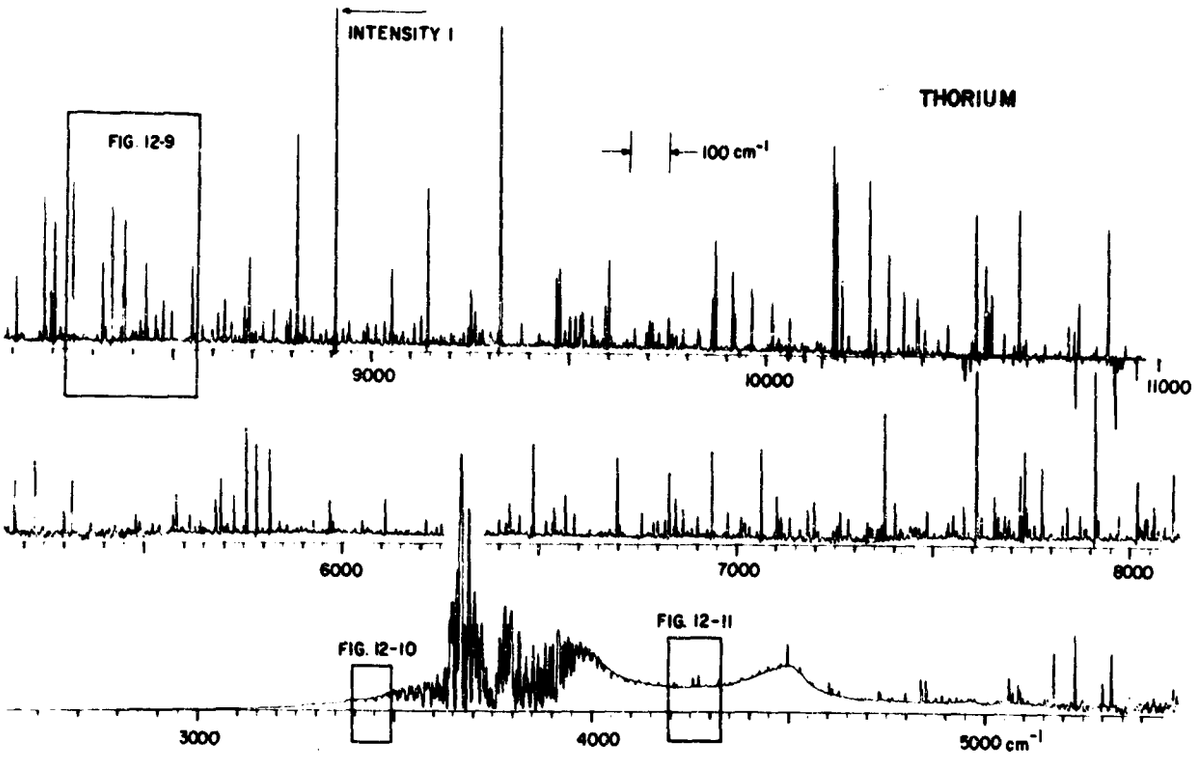


Figure 12-8. Low Resolution Emission Spectrum of Thorium Computed From the First 10^4 Samples of a 10^6 Samples Interferogram

time was 10 hr for 10^6 samples. Had the system been run at maximum speed, this could have reduced to about 6 hr. Resolution is 20×10^{-3} cm^{-1}.

Figure 12-8 represents the whole spectrum of thorium computed from only the first 10^4 points of the million samples interferogram. The corresponding resolution limit is 2 cm^{-1} and the recording time 6 min; the spectrum is normalized to the intensity of the strongest line at 8904 cm^{-1}.

Between 9500 cm^{-1} and 11,000 cm^{-1} a few negative lines coming from the second order can be seen. This means that the spectral range actually recorded extends down to 12,500 cm^{-1}. At about 4000 cm^{-1}, a continuous background emitted by the hot quartz

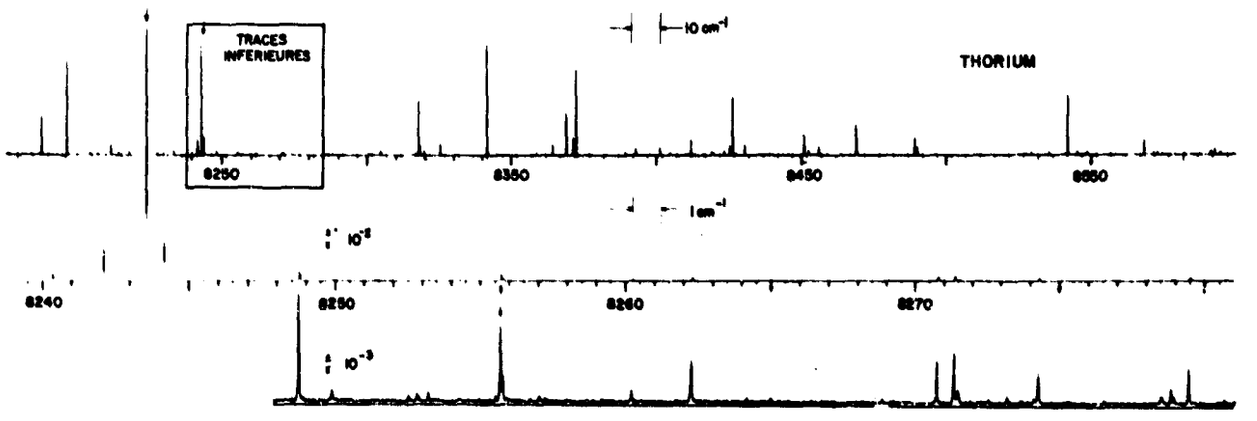


Figure 12-9. The Upper Trace is a Part of the Spectrum Plotted on Figure 12-8 but Computed From 10^8 Samples. The two other traces represent a part of the spectrum with the maximum resolution ($\delta\sigma = 20\times10^{-3}$ cm^{-1})

tube appears. The atmospheric water vapor explains the absorption around 3800 cm^{-1}.

The upper trace of Figure 12-9 shows the part of the spectrum which is singled out on Figure 12-8 at about 8300 cm^{-1}. The number of computed points, 10^5, corresponds to a 0.2 cm^{-1} resolution and the recording time is now 1 hour. The middle trace represents 40 cm^{-1} of the spectrum computed from 10^6 samples ($d\sigma = 20 \times 10^{-3}$ cm^{-1}); the intensity scale has been expanded by 10. On the lower trace, two spectra obtained from two different interferograms are drawn. The intensity scale has been increased again so that noise is perceptible. Figures 12-10 and 12-11 show two other portions of the spectrum with the maximum resolution.

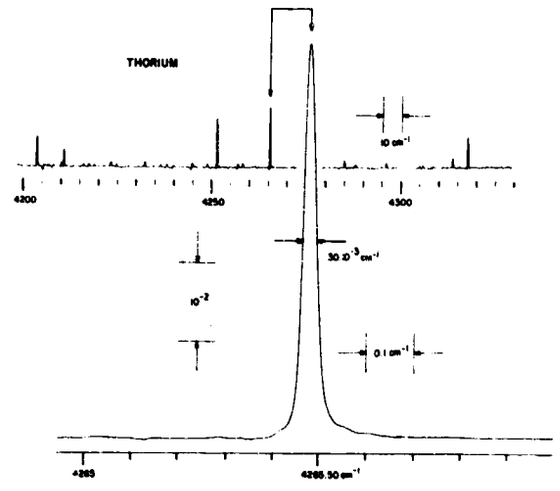


Figure 12-10. Upper Trace: Spectrum Computed From 10^5 Samples; Lower Trace: Spectrum Computed from 10^6 Samples

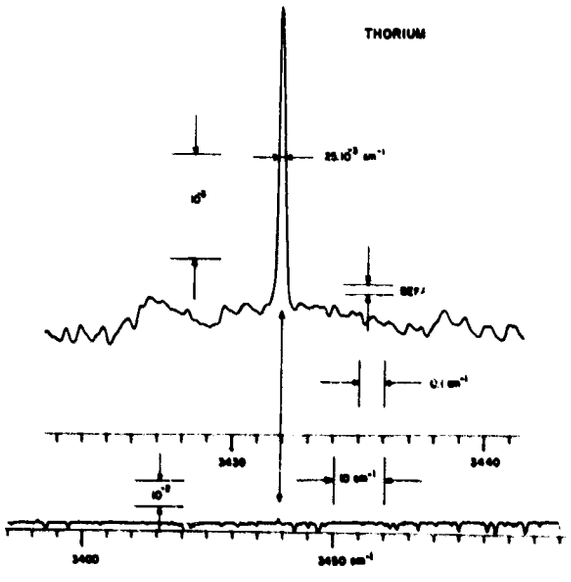


Figure 12-11. Upper Trace: Spectrum Computed From 10^6 Samples; Lower Trace: Spectrum Computed From 10^5 Samples

The sources are unstable and require many precautions. The frequency of the internal modulation and the sampling rate fall within the frequency range of their oscillations. This explains why the receiver noise has not been reached yet. However, in the case of the thorium spectrum, one obtains $Q = 10^7$ and a signal-to-noise ratio for the more intense lines equal to 12,500.

Figure 12-12 is another example of an emission spectrum with an holmium source. All the recording conditions are the same as for thorium. The lines present the characteristic hyperfine structure of Ho ($S = \frac{7}{2}$). The quality factor is equal to 2.5×10^7; when compared with the planetary spectra obtained by Connes et al.[3,4]; this is an improvement by a factor of 10. Figure 12-13 compares the weakest structure recorded by Verges with a SISAM (recording time 20 min for 5 cm^{-1}), with the same structure obtained by the computation of the 10^6 samples interferogram. This structure is 25 times more intense than the weakest that can be detected on our spectra.

12–5.2 Absorption Spectra

Because of the lack of a long cell, highly absorbing gases (N_2O, ICH_3) have been the main concern of this study.

The white light source was an iodine quartz tungsten lamp. The white light beam is reflected once at the end of the 5-m long cell and passes through 10 m of gas. Since all the apparatus is not yet enclosed in a vacuum tank, some information in the regions of the water bands is lost. In this case also, one is not yet limited by the receiver noise but by the "light noise" due mostly to small vibrations of the long path cell mirror mounts. Since the absorption lines of N_2O and ICH_3 gas at low pressure actually show the room temperature Doppler width (5 to 10×10^{-3} cm^{-1}), the instrumental resolution of 5×10^{-3} cm^{-1} is needed. To cover the full 3000 to 10,000 cm^{-1} region, 4×10^6 samples should be needed. Up to now, not more than 10^6 samples transformed could be used and the range was limited to about 2000 cm^{-1}. Considering the latest reductions in computing time (Delouis and Connes at the Aspen Conference), one should no longer be afraid of transforming 4×10^6 samples.

Figure 12-14 gives the whole low-resolution spectrum of N_2O at 2 mm Hg pressure between 3800 cm^{-1} and 5600 cm^{-1}. The interferogram from which the computation has been done is the first part of an interferogram including 450,000 samples recorded during 5 hr 30 min (length of the steps $2.695\,\mu = 77 \times 350$ Å, peak-to-peak amplitude of internal modulation $1.05\,\mu = 30 \times 350$ Å). For the spectrum in Figure 12-14, only 4500 samples were used. This corresponds to a 4 min recording time and a resolution limit equal to 0.8 cm^{-1}. Consider the $20°1 \leftarrow 00°0$

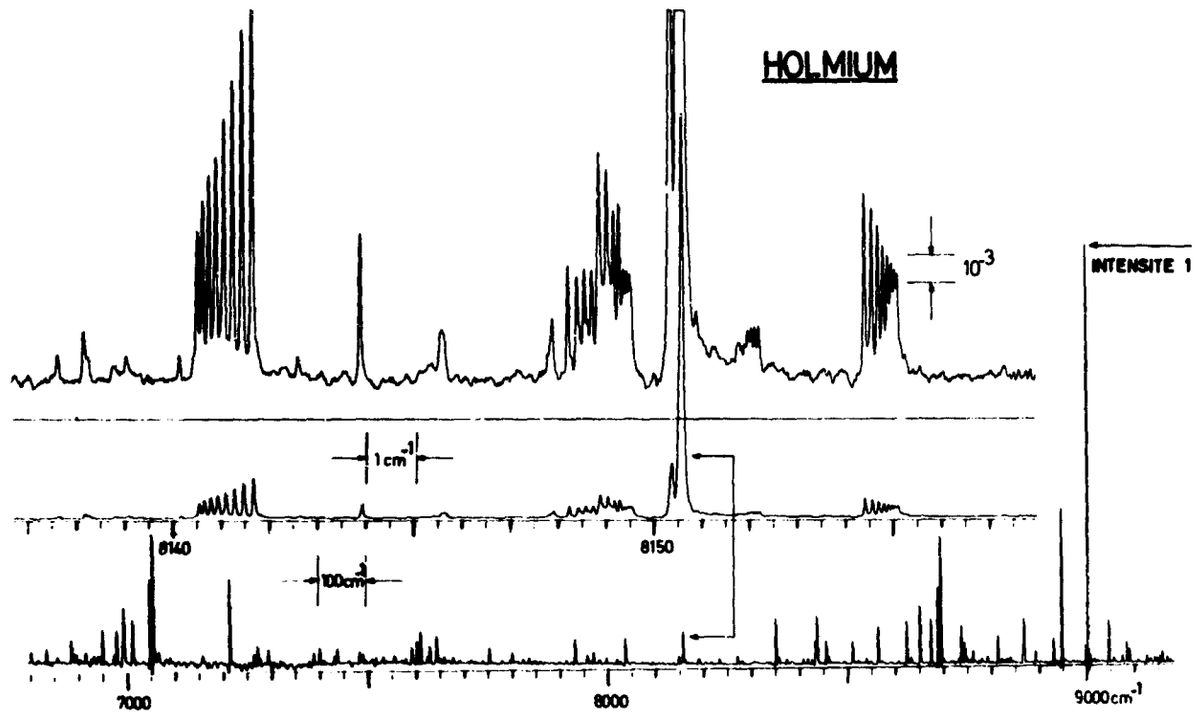


Figure 12-12. Emission Spectrum of Holmium. The two upper traces are computed with maximum resolution. (10^6 samples; $\delta\sigma = 20 \times 10^{-3}$ cm^{-1})

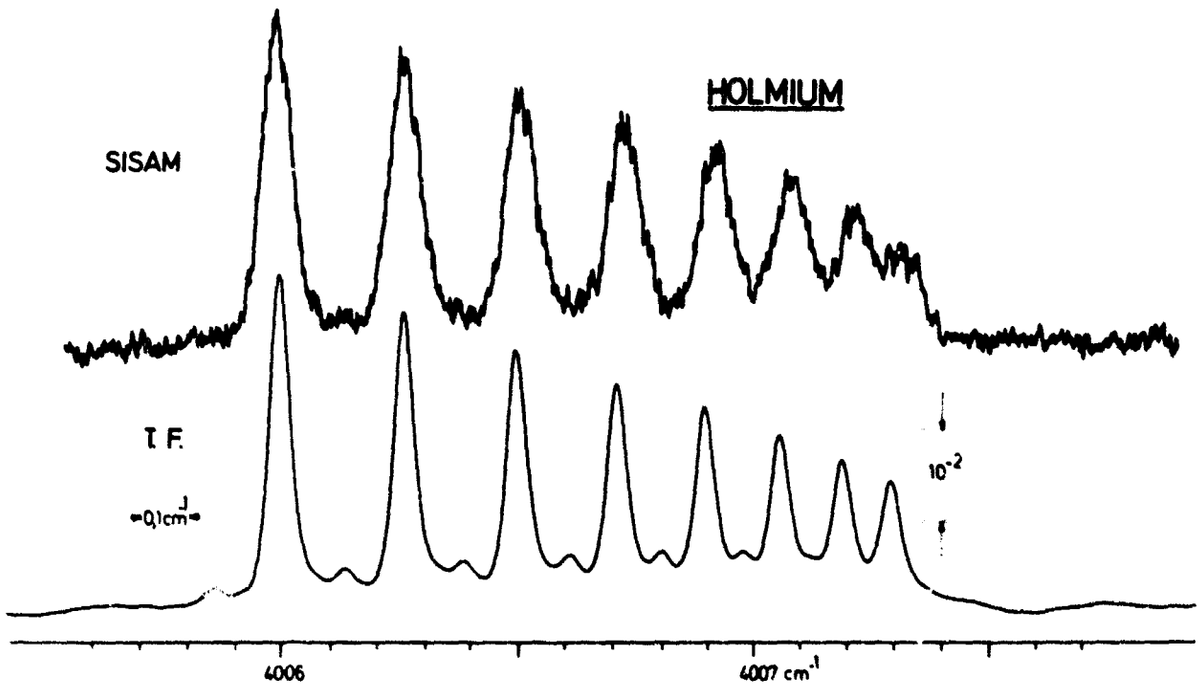


Figure 12-13. Comparison of the Same Feature Recorded by Scanning Spectrometer (SISAM) and Fourier Spectrometer

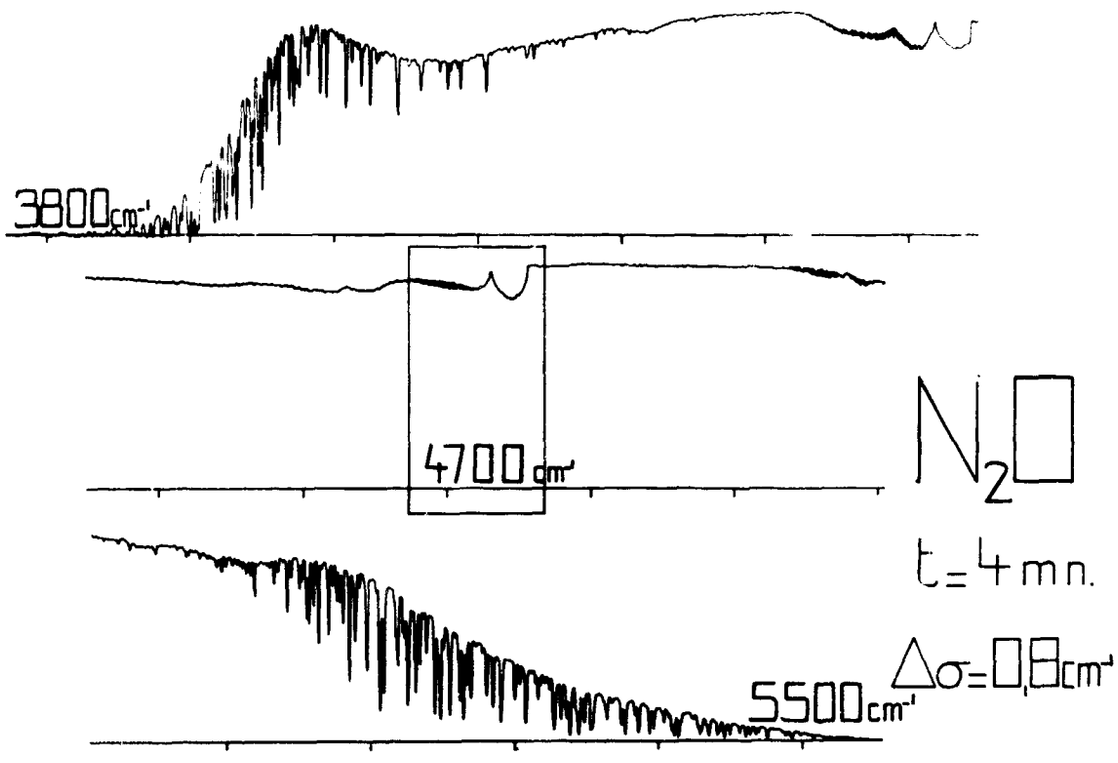


Figure 12-14. Low Resolution Absorption Spectrum (N_2O) From About 5000 Samples

$(2\nu_1+\nu_3)$ band at 4730 cm^{-1} inserted in the figure. It is plotted again on the lower trace of Figure 12-15. The middle trace has been calculated from 45,000 samples ($d\sigma = 90\times10^{-3}$ cm^{-1}, recording time 45 min). The rotational structure of the $20^01 \leftarrow 20^00$ band is well resolved. The hot band, $21^11 \leftarrow 01^10$, begins to appear but the resolution is not yet good enough to show it in full. The upper trace is the plot of a 10 cm^{-1} region close to the band center. The very high resolution (8×10^{-3} cm^{-1}) allows complete separation of the two bands and observation of the l-type separation (increasing with the value of J) which occurs in the $21^11 \leftarrow 01^10$ band. Two spectra computed from two different interferograms are superimposed and give an idea of the reproducibility. The whole spectrum with this wavenumber scale (0.15 cm^{-1}/cm) is 120 m long. The quality factor Q is equal to 5×10^7. The signal-to-noise ratio is 200.

Another example of a very high resolution absorption spectrum is given in Figure 12-16 with ICH_3 (gas pressure $\simeq$2 mm Hg). This spectrum covers 600 cm^{-1}, from 2800 cm^{-1} up to 3400 cm^{-1} (2×10^6 samples; length of the steps: $9\,\mu = 255\times350$ Å; peak-to-peak amplitude of internal modulation: $1.5\,\mu = 43\times350$ Å, recording time 5 hr). The limit of resolution is the smallest obtainable with the interferometer ($d\sigma = 5\times10^{-3}$ cm^{-1} with $\Delta_{max} = 2$ m). The Doppler width of the lines is equal to 4×10^{-3} cm^{-1}. For this spectrum, the quality factor is equal to 3×10^7.

12-6 CONCLUSION

The interferometer has given the kind of results hoped for, both in resolution and wide spectral range coverage.

Compared to previous generations of Fourier interferometers,[1,3] the sampling speed has been increased from about 5 samples/sec to 50 samples/sec and the diameter of the beam has been increased from 2.5 cm to 8 cm. The reliability has also been much improved, which was essential for increasing the number of samples from about 5×10^4 to 10^6, or more.

The improvement compared to previously available techniques for recording near infrared spectra is striking in the case of the rare earth emission spectra; this resolution could only be realized from a large number of separate F.P. or SISAM recordings. This could only be made after a preliminary low-resolution study of the entire spectrum.[5,6,7] The whole program took several months. The accuracy with which line positions can be measured from high-resolution Fourier interferometers has been demonstrated by Pinard.[1]

Now the problem of accurate absolute measurement of wavenumbers has to be solved. The interferometer must be put in the vacuum and the absolute wavelength of the 3.50 μ Xe line from the superradiant tube must be measured by recording the spectrum of the standard krypton lamp.

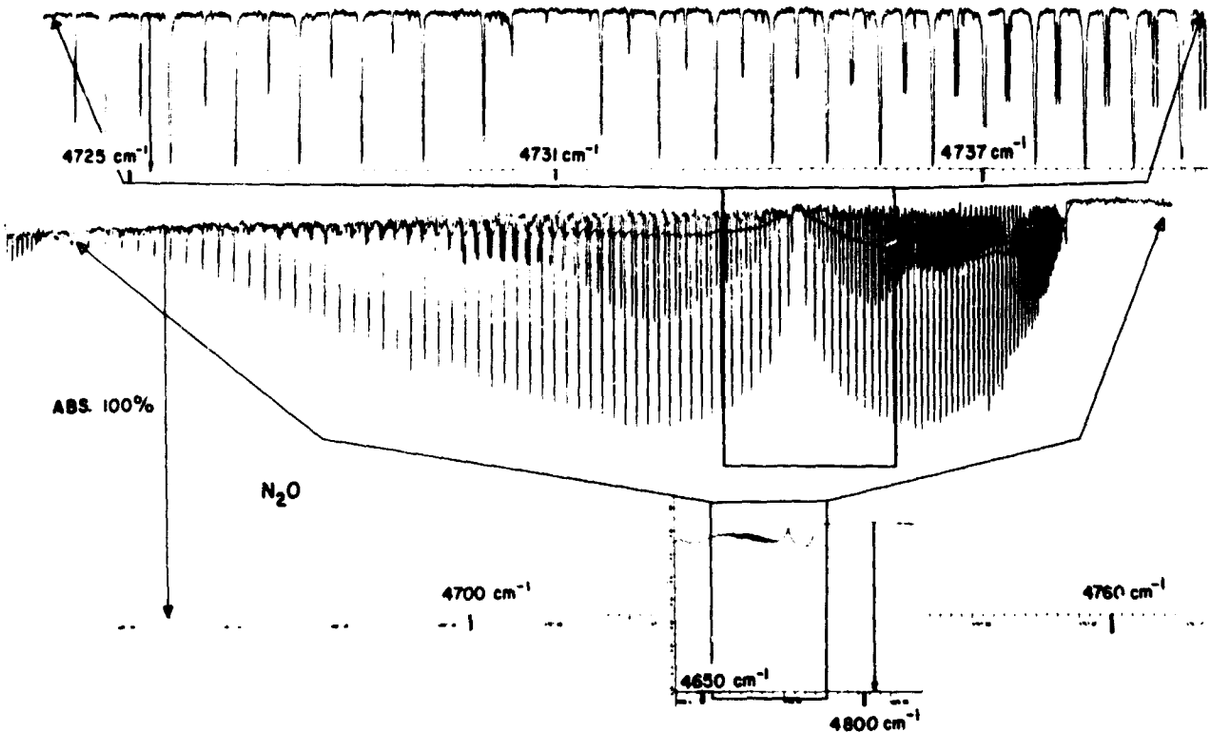


Figure 12-15. N_2O From the Same Interferogram But With Different Resolutions (Maximum Resolution = 8×10^{-3} cm^{-1})

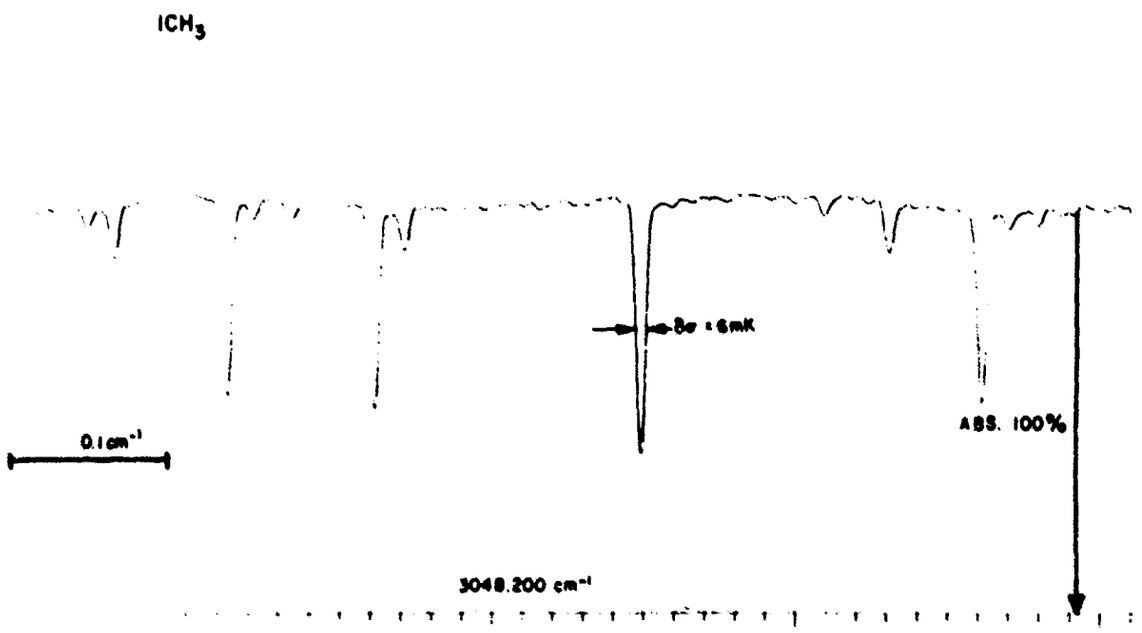


Figure 12-16. Very High Resolution Spectrum of ICH_3 ($\delta\sigma = 5 \times 10^{-3}$ cm^{-1})

References

1. Pinard, J. (1967) *J. Phys.* 28, C2: 136.

2. Connes, J., Connes, P., Delouis, H., Guelachvili, G., Maillard, J. P., and Michel, G. (1970) *Nouvelle Rev. Opt. Appl.* No. 1.

3. Connes, J., Connes, P., and Maillard, J. P. (1967) *J. Phys.* 28, C2: 120.

4. Connes, J., Connes, P., and Maillard, J. P. (1969) Atlas des spectres dans le proche infrarouge de Venus, Mars, Jupiter et Saturne, *Ed. du CNRS.*

5. Verges, J. (1969) *Spectrochim. Act.* **24B**: 177.

6. Camus, P., Guelachvili, G., and Verges, J. (1969) *Spectrochim. Act.* **24B**: 373.

7. Blaise, J., Marillon, C., Schweighofer, M. G., and Verges, J. (1969) *Spectrochim. Act.* **24B**: 405.

Contents

13. A Method of High Resolution Fourier Transform Spectroscopy

H. L. Buijs
Université Laval
Quebec, Canada

Abstract

A tilt-compensated two-beam interferometer with a continuously variable path difference of 100 cm was used to study molecular constants and the phenomenon of collision narrowing in the static-field-induced absorption in hydrogen gas near 2.5 μ. Line widths down to 0.01 cm^{-1} were measured and frequencies were calculated to a precision of 0.0005 cm^{-1}. The design of the interferometer and the methods of sampling and numerical filtering are discussed.

13–1 INTRODUCTION

In order to study sharp absorption features in the infrared, such as the static-field-induced absorption lines in hydrogen gas, it appears desirable to apply the Fourier transform method of spectroscopy not only because of the multiplex advantage, but also because of superior frequency and intensity calibration and because the instrumental profile is well known. On the other hand, until recently the application of the Fourier transform method to high resolution work has been restricted by the requirement of a stable interferometer for long path-difference displacements and the need for computational techniques capable of transforming records containing very many spectral elements.

13–2 DESCRIPTION

The interferometer used by the author has been described in the literature before[1] and is constructed in such a manner that the interference pattern at the exit is unaffected by small changes in orientation of the moving mirror elements (see Figure 13-1). This particular design of interferometer appears to be unique and has resulted in an instrument that is very compact and simple to maintain. The moving mirror elements can be displaced 25 cm for a maximum path-difference change of 100 cm. The path difference is

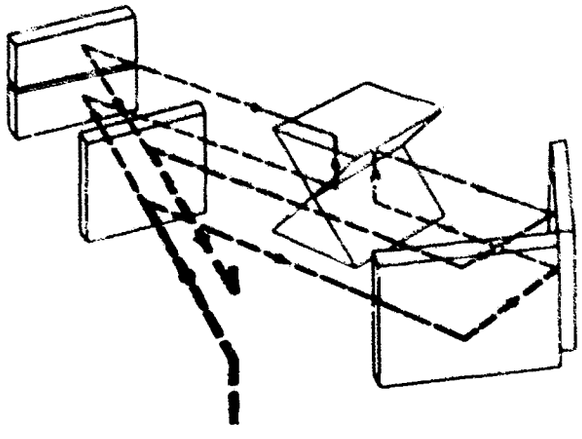


Figure 13-1. Schematic Diagram of Interferometer Showing the Path of the Light Through the Interferometer. The fringes are not affected by small rotations of the reflecting prisms. The beam splitter is 10 cm × 15 cm calcium fluorite

monitored by means of a stabilized single mode He-Ne laser producing a reference interferogram which is also used to indicate the state of alignment of the instrument. The path difference is varied continuously by means of a simple mechanical drive.

During the uniform displacement of the moving mirrors, the interferogram is sampled nearly instantaneously at intervals determined from the reference interferogram. In this continuous scanning method, the digital information becomes representative of the total amount of information available during the observation time when it is possible to reproduce precisely the analog signal from the digital information. This is simply accomplished when the total frequency content of the analog signal is limited, via some electrical filter network, to a finite bandwidth, and when the sample frequency is at least twice this bandwidth.

In many cases this means that the interferogram must be oversampled with respect to the relevant signal in order to prevent overlapping of noise features. Compared with the better known method of start-stop motion, it appears that this mechanical simplification results in a larger data handling problem. It has been found, however, that with the continuing rapid development of high speed electronic computer facilities, the handling of large amounts of data is neither difficult nor expensive.

The field induced fundamental rotation-vibration spectrum of hydrogen is centered about 2.4 μ in the infrared and consists of a number of widely spaced sharp lines in the region from 2 to 2.5 μ. We have constructed an optical frequency synthesizer consisting of a double-pass reverse dispersion grating spectrograph with which a number of bands out of the continuum of an infrared source can be selected (see Figure 13-2).

By selecting narrow bands of radiation at the frequencies at which absorption lines occur, much unnecessary radiation is prevented from entering the interferometer and hence the dynamic range problem is not unnecessarily complicated. The interferogram was sampled once for every two fringes of the laser reference, which corresponds to a sample interval of about 1.25×10^{-4} cm. This sample interval is not small enough to prevent some overlapping of noise, and a corresponding sacrifice in signal-to-noise ratio is anticipated. A typical interferogram consists of 800,000 measurements and is recorded on magnetic tape by means of a properly buffered digital tape drive.

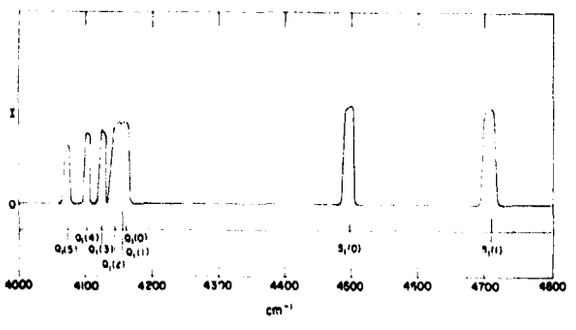


Figure 13-2. Spectrum of Bands Selected by Optical Frequency Synthesizer. The positions at which absorption occurs in the hydrogen gas are indicated with a stick spectrum below

It is not difficult to perform a complex Fourier transform on the previously mentioned interferogram, making use of a fast Fourier transform algorithm and a factoring technique developed by the author[3] which allows efficient use of non-random-access bulk storage. However, the spectral range that can be covered with these interferograms contains many gaps at which no useful information occurs; hence, it was found more efficient to carry out some numerical filtering and consequent reduction in data points. The numerical filtering is done in two steps:

1) When the laboratory-generated tape is interpreted on a large computer, the data are convolved with a broad Gaussian filter function covering just the spectral region of interest. This allows a reduction in points of a factor of 4, and the newly generated interferogram, containing 200,000 points, is stored in compressed form on an inventory tape. During this first operation on the interferogram the exact zero path difference location is computed by means of a low resolution complex transform.

2) Prior to transformation, the new interferogram is convolved with a composite filter function having a multiband response. During this convolution the data are also interpolated to provide a point at zero path difference. The reduction in data in this second step is about a factor of 2.5, and the different parts of the spectrum are no longer related by a linear frequency scale. The different bands of interest are extracted from different aliases in this second filter operation, so that the bands occur adjacent to one another in the principal alias and hence occupy the smallest possible frequency band (see Figure 13-3). Finally, a Fourier transform is performed on a single-ended interferogram consisting of about 90,000 data points. The time of computation on our 360/67 computer is about 5 min for tape reading and the first filter operation, and about 6 min for the second filter operation, interpolation, and transformation.

The study recently completed by means of the described technique involved the determination of line profiles and central frequencies as a function of pressure in the field-induced absorption spectrum of hydrogen gas at room temperature. These spectra constitute the first quantitative observations of the phenomenon called "collision narrowing" in the infrared. We have obtained frequencies extrapolated to zero density with an estimated accuracy of 0.0005 cm^{-1}, and it was possible to resolve lines as narrow as 0.01 cm^{-1} (see Figure 13-4). The same work performed with a highly developed grating spectrograph has failed to yield any quantitative measurements of the collision narrowing phenomena, and has fallen short of the frequency calibration accuracy reported in our work.[4]

The work is presently being extended to a study of the HD molecule for which we hope to measure the matrix elements of a very weak allowed dipole moment, in addition to the study of the field-induced dipole spectrum.

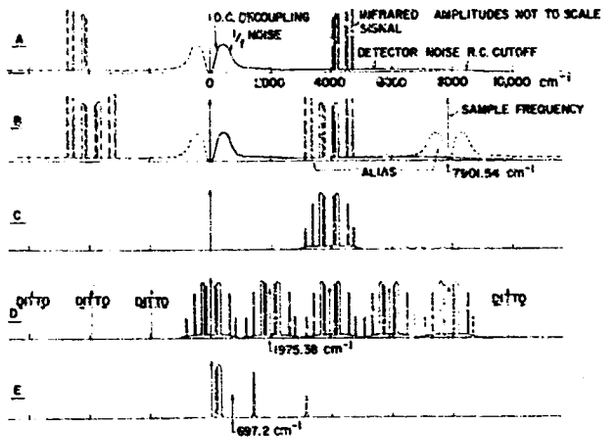


Figure 13-3. Composite Filter Function with Multiband Response. a) Spectral contents of interferogram prior to digitizing. b) Spectral contents of interferogram in digitized form when represented by 800,000 data points. c) Effect of first numerical filtering operation when the resulting interferogram is highly oversampled. d) Spectral contents of 200,000 point interferogram. e) Effect of second numerical filtering operation when the resulting interferogram is highly oversampled. The final sample frequency possible is also indicated.

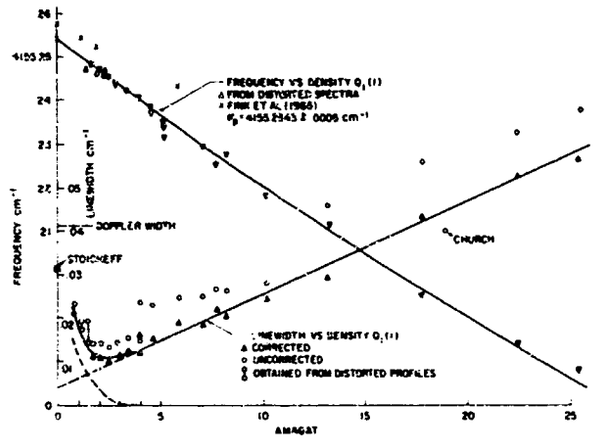


Figure 13-4. Variation of Central Frequency and Line Width as a Function of Density for a Typical Spectral Line in Hydrogen. For the line width the open circles are actual observed width and the triangular dots are deduced width based on comparing actual data with synthetic spectra generated with a computer

References

1. Buijs, H. L., and Gush, H. P. (1966) *Méthodes Nouvelles de Spectroscopie Instrumentale*, Orsay.

2. Buijs, H. L., and Gush, H. P. (1967) *J. Phys.* **28**: C2.

3. Buijs, H. L. (1969) *Appl. Opt.* **8**: 211.

4. Fink, U., Wiggins, T. A., and Rank, D. H. (1965) *J. Mol. Spectry.* **18**: 384.

Discussion

Comment (Uwe Fink, University of Arizona): I noticed that there was a small systematic difference of 0.003 cm^{-1} between your frequency measurements and the ones of Fink, Rank, and Wiggins. A measurement of less than about 0.020 cm^{-1} for the halfwidth of the S(1) line of the fundamental in the H_2 quadrupole spectra was done about 1964 by D. H. Rank, providing the first quantitative measurement of the collision narrowing.

Q. (A. L. Fymat, Jet Propulsion Laboratory): The phenomenon of collision narrowing of spectral lines is of interest in the investigation of dense gases such as, for example, the massive atmospheres of the major planets in our solar system (Jupiter, Saturn, . . .) I should like to know if Dr. Buijs has applied his measuring technique and results to this case.

A. (C. Cumming): No.

Comment (Uwe Fink): The problem of the hydrogen quadrupole lines in Jupiter, including collision narrowing, has been systematically investigated in a paper by Uwe Fink and M. J. Belfin in *J. Am. Sci.*, Sept. 1969.

Contents

14. A High Resolution Far Infrared Interferometer

R. B. Sanderson and H. E. Scott
Department of Physics
The Ohio State University
Columbus, Ohio 43210

Abstract

A high resolution interferometer in the Michelson configuration has been designed and constructed for the purpose of studying the optical properties of gases in either the symmetric or asymmetric mode of operation. The performance of this instrument in the 20 to 250 cm^{-1} spectral region is demonstrated by measurements on several simple molecules—CO, HCl, and H_2O. A working resolution of 0.05 cm^{-1} is verified by observation of the isotopic splitting in HCl and the resolution of several doublets in the H_2O spectrum. Measurement of the self-broadened half widths of the pure rotational CO transitions, $J=7$ through 20, is also discussed.

14–1 INTRODUCTION

A high resolution, far infrared interferometer has been constructed for the purpose of studying the optical properties of gases by Fourier transform techniques.[1] The primary interest is in determination of transition probabilities and in the spectral effects of intermolecular interactions. For measurements of this type, the instrument should have the following capabilities. Since the spectral widths of molecular transitions are typically of the order of 0.1 cm^{-1}, the resolution should be at least this high. It should also be possible to use either symmetric or asymmetric modes. The usefulness of the asymmetric mode has been demonstrated by previous work in this laboratory[2,3] and by Gebbie and Chamberlain.[4,5] Finally, provision should be made for future extension to near infrared operation. In this region the instrumental problems are more difficult, and more sophisticated methods for determining the optical path difference are required.

In order to meet these requirements, the interferometer is built in the Michelson configuration. The radiation in the arms must be well collimated, and the arms must be long enough to accommodate a cell of reasonable length in the asymmetric position. Such a configuration has, in general, poor throughput, unless the mirrors are inconveniently large and it is difficult to align. In the present instrument these problems have been minimized by constructing the optical system on axis and using "cat's-eye" retroreflectors[6,7] in the arms. The use of retroreflectors greatly simplifies alignment and reduces the effects of yaw and pitch in the motion of the movable mirror assembly. Since

the secondary mirror of the cat's-eye assembly obstructs the beam center, further loss due to the on-axis operation is not serious.

The general plan of the interferometer is shown in Figure 14-1. It is built in two vacuum tanks salvaged from an old spectrometer. These tanks are mounted on tripod supports and fastened together rigidly at

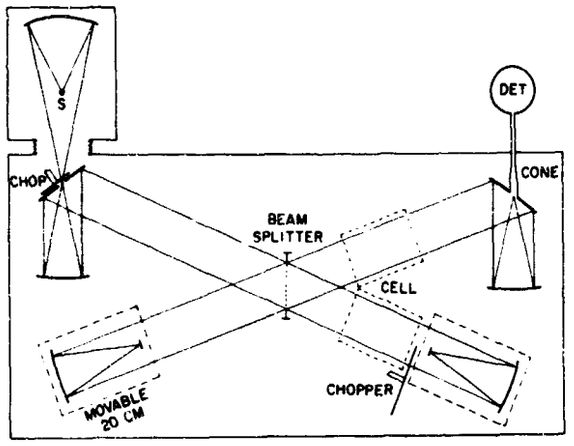


Figure 14-1. Optical Diagram of Interferometer

the bases. The bed plates are mounted on three point supports directly over the tank legs. This arrangement permits the tank walls and base to flex when the system is evacuated without misaligning the optics. The acute angle between the arms of the interferometer was dictated by the size of the available tank. The system can be evacuated to a pressure of 1 μ by a large mechanical pump. The entire interferometer is maintained a few degrees above ambient by a heating belt around the outside of the tank. This improves the thermal stability of the interferometer.

14-2 OPTICS

The source is a Phillips HPK 125 W high pressure mercury arc. It is imaged at a 2 cm hole in the entrance flat. Cooling is provided by water flowing through copper tubing wrapped around the base of the socket and around the portion of the optical path within the source tank. The water line also provides cooling for motors used elsewhere in the interferometer.

The collimating mirror, telescope mirror, and cat's-eye primaries are all identical paraboloids. They have a diameter of 15 cm and a focal length of 25 cm. The primary of the cat's-eye forms an image of the source on the small convex secondary mirror. The curvature of the secondary is chosen so that the beam splitter is imaged back upon itself by the cat's-eye assembly. This arrangement optimizes the throughput without causing appreciable loss of resolution. Use of the convex secondary improves the signal by a factor of two over that obtained using a plane secondary.

The exit optics image the source at the hole in the exit flat. A 6 deg brass condensing cone reduces the beam diameter to 1 cm. The radiation then goes through a brass light pipe to the detector in a cryostat. The detector is a Zn doped GaAs bolometer* maintained at 1.5° K.

The use of cat's-eyes greatly facilitates alignment. The interferometer arms are about 70 cm long. With plane end mirrors, the adjustment of these mirrors and of the beam splitter would be very critical. With the cat's-eyes, alignment is done by the following simple procedure. Cross hairs are placed in the hole in the entrance flat. When the cross hairs are simultaneously imaged on both secondaries and at the hole in the exit flat, it is possible to observe interference fringes. Final alignment is done with the cat's-eye mirrors. An important advantage of this arrangement is the insensitivity of the alignment to the orientation of the beam splitter. Beam splitters can thus be changed without realigning the optics.

14-3 MECHANICAL

A critical part of the interferometer is the mechanism which must translate the movable mirror assembly in precisely equal steps. In the far infrared, the tolerances are greater than at shorter wavelengths, and a simple system is satisfactory. The movable cat's-eye is mounted on a 20 cm comparator screw.† This permits, in principle, a resolution of 0.025 cm^{-1} The screw is driven by a stepping motor** geared to change the optical path by 0.6333 μ/step. The stepping motor is activated by a preset indexer** which may be set to advance the stepping motor the appropriate number of steps whenever it receives the triggering signal. The step size is arbitrary in the present operation, but it was chosen to facilitate future use of a He-Ne laser control system at higher frequencies.

The signal is modulated at 10 Hz by one of two choppers. The first chopper, located at the entrance flat, is used only for alignment. In normal operation, the radiation is modulated by a chopper in the fixed arm. This chopper is a well polished aluminum disc which extends only half way across the beam and is aligned so that the intensity reflected from the blade is nearly equal to the intensity transmitted by the cat's-eye in the absence of the blade. The reflected radiation is effectively incoherent with the radiation transmitted by the cat's-eye, and thus compensates the constant term in the detector signal without affecting the interferogram. This arrangement appreciably reduces the dynamic range requirements of the detector system. The chopper also modulates the light falling on a photodiode from a small bulb to provide a reference signal.

* The detector was furnished to us by R. G. Wheeler of Yale University.
† D. W. Mann, Co., Lincoln, Mass.
** Superior Electric Co., Bristol, Conn.

14-4 ELECTRONICS AND DATA HANDLING

The output from the detector is amplified by a Tektronix model 122 preamplifier and a Perkin-Elmer model 107 amplifier. The amplified signal is read by a Hewlett-Packard model DY-2401B integrating digital voltmeter. The voltmeter has been modified so that the integration time can be varied between 1 and 16 sec. When the integration is completed, the voltmeter produces a signal which activates a card punch and triggers the preset indexer to advance the movable cat's-eye. The timing between steps is provided by the internal circuitry of the voltmeter. The actual transformation is done on the IBM 360-75 at the Ohio State University Computer Center.

14-5 SAMPLE

The sample cell is normally placed in one of the two locations indicated in Figure 14-1. The cell mount is pivoted under the beam splitter so that the cell can be rotated into the symmetric or asymmetric position without breaking either the instrument vacuum or the connection from the cell to the sample-handling system. This arrangement permits obtaining both symmetric and asymmetric data on the same sample. It is, however, necessary to evacuate the cell to obtain a background.

The cell which has been used for all measurements to date is 22.5 cm long and has a diameter of 11.2 cm. The windows are polypropylene, about 1.5 mm thick. They were wedged by cutting one surface to a conical form on a lathe. One window is convex; the other is concave. This makes the optical path roughly the same at all points of the aperture. While only slightly reducing the effectiveness of the wedging in reducing the channel spectra in the symmetric mode, this arrangement reduces the "pseudo-coherence" effect[8] which would normally make wedged windows unusable in the asymmetric mode.

14-6 OPERATION

The spectral region is determined by the choice of Yoshinaga filter and beam splitter. The most commonly used combinations are listed in Table 14-1. In addition, black polyethylene is always used at the exit hole, and there is a sapphire window on the detector housing.

The presently used Fourier transform programs can accept a maximum of 8192 input points in the symmetric mode. The computer adds an equal number of zeros on the negative side of the central peak, so that the transformation actually extends up to 16,384 points. The computer is large enough for this to be doubled to 32,768 points. However, the present restriction limits the resolution only at wavenumbers above 200 cm^{-1}. The practical limit to the number of points is the speed of the data acquisition system. For a very large number of points, the total running time becomes so long that it is difficult to maintain instrumental stability. The normal maximum elapsed time for a run is 6 hr. For runs which would require more time, it is better to reduce the spectral band by filtering out the low frequencies. This permits increasing the step size and thus reduces the number of points without sacrificing resolution. The integration time can be increased and more efficient use made of the running time.

Table 14-1. Beam Splitter and Filter Combinations Frequently Used

Optimum Spectral Region (cm^{-1})	Beam Splitter	Yoshinaga Filter	Step Size (μ)
25– 70	100μ metal mesh	KBr, KCl, BaF_2	44
50–130	50μ metal mesh	BaF_2	25
100–240	6.4μ (thickness) mylar	BeO, LiF	19

14-7 PERFORMANCE

Performance will be illustrated by presenting the results of several measurements on CO, HCl, and H_2O.

14-7.1 Wavenumber Precision

The wavenumber scale calculated in the Fourier transform program is only approximately correct. In this program the optical path difference between sampling points is read as an integral multiple of 0.6333μ. This is a nominal step size based on the gear ratio and rounded to four figures. In addition there is an error due to the divergence of the beam. There would be a fractional correction to the observed wavenumbers of -4×10^{-4} if the input solid angle were uniformly filled.[1] Since the solid angle is not uniformly filled, the correction is less than that. To determine wavenumbers to better than a few hundredths of a reciprocal centimeter, the calibration must be experimentally determined. Since both the above effects produce an error proportional to wavenumber, the deviations can be represented by the form

$$\sigma_{obs}-\sigma_{calc}=\alpha\sigma_{calc}. \qquad (14\text{-}1)$$

The correction is specified by a single number α. It is most conveniently taken into account by defining

an effective step size in the transform program

$$\Delta X_{eff} = \Delta X_{nom}(1+\alpha). \qquad (14\text{-}2)$$

Carbon monoxide is used as the calibration standard. Below 100 cm^{-1} this molecule has a convenient series of narrow, closely spaced lines whose wavenumbers[9] are known to better than 0.001 cm^{-1}. By interpolating the spectrum[10] as it comes from the computer with points spaced about 0.002 cm^{-1} apart, a symmetrical line shape is obtained, and its center can be located to about 0.001 cm^{-1}.

The observed wavenumbers vary with filtering, step size, and resolution. However, for any given operating conditions, the consistency is much better. The values of α range from nearly zero to about -2×10^{-4}. These values are of the order of magnitude expected for the uncertainty in the step size and in the beam divergence. With this correction taken into account, the maximum deviation for lines with good signal-to-noise ratio is 0.002 cm^{-1}, or about one twenty fifth the resolution width of the instrument.

14-7.2 Resolution

To permit operation in the asymmetric mode, the white light position is set at the center of the drive screw. This choice limits the maximum optical path difference to 20 cm and the resolution to 0.05 cm^{-1}. The resolution actually achieved is demonstrated by observation of the splitting of several doublets in HCl and in H_2O.

Figure 14-2 shows two regions of the H_2O rotational spectrum which have often been used to illustrate the performance of far infrared spectrometers. The quoted wavenumbers have been determined by Rao and Williamson from measurements on the three fundamental vibrational bands.[11] In the 38 cm^{-1} region, the weak line in the center of the triplet is separated from the strong line at 38.786 cm^{-1} by 0.155 cm^{-1}. The ratio of the strengths is about 75:1. The 59.9 cm^{-1} doublet has a spacing of 0.093 cm^{-1}. It is almost completely resolved.

Figure 14-3 shows the region around 100 cm^{-1} which includes two closely spaced doublets. The pair

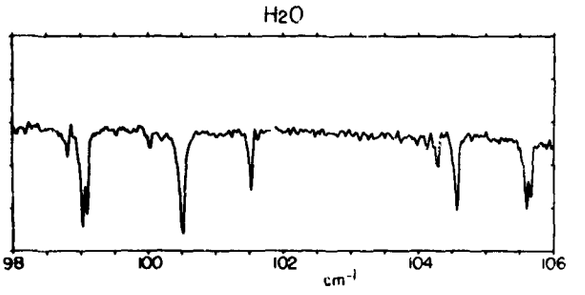


Figure 14-3. Water Vapor Spectrum. The maximum optical path difference was 17.79 cm. There were 7024 input points, each integrated for 1 sec. The pressure was 2 Torr

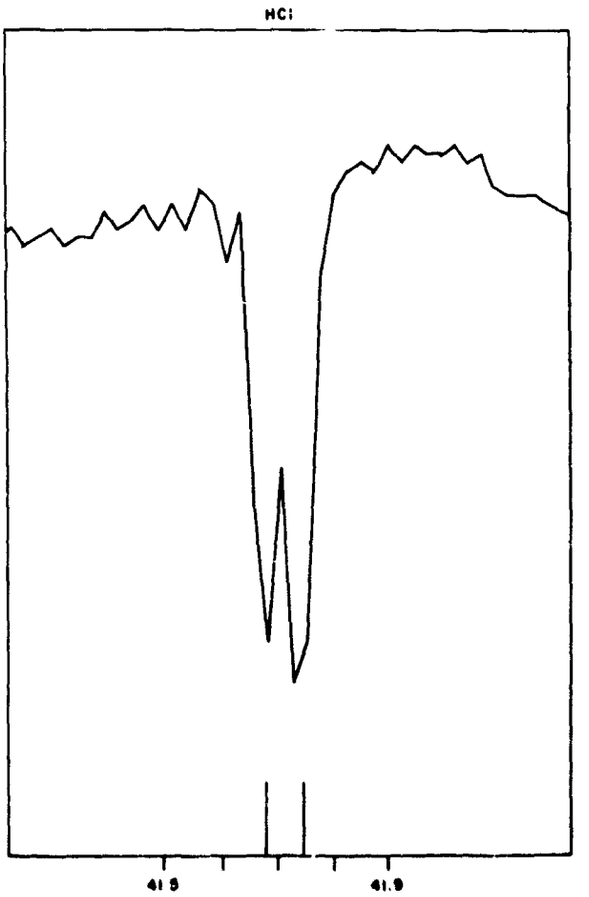


Figure 14-4. Hydrogen Chloride Spectrum. The maximum optical path difference was 17.18 cm. There were 3392 input points, each integrated for 4 sec. The pressure was 4 Torr

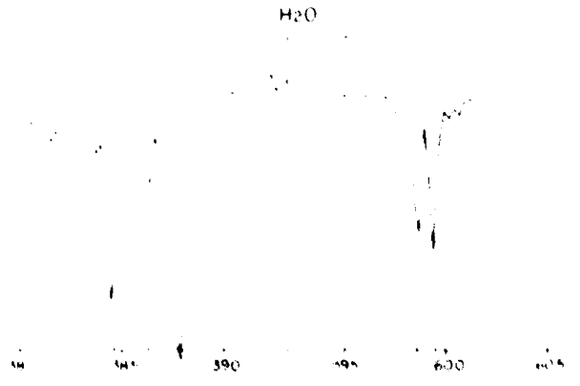


Figure 14-2. Water Vapor Spectra. The maximum optical path difference was 17.18 cm. There were 3392 input points, each integrated for 4 sec. In the 38 cm^{-1} region the pressure was 8 Torr; in the 59.9 cm^{-1} region it was 4 Torr

of lines at 99 cm^{-1} has a spacing of 0.082 cm^{-1}, and the separation of the 105.6 cm^{-1} doublet is 0.052 cm^{-1}.

The rotational lines of HCl are all double due to the presence of the two chlorine isotopes. This doubling is given approximately by the formula

$$\sigma(H^{35}Cl)-\sigma(H^{37}Cl)=0.032(J+1),$$

where J is the lower state rotational quantum number.[12,13] In Figure 14-4, the splitting of the $J=1$ transition at 41 cm^{-1} is shown. The separation here is 0.0638 cm^{-1}.

The observed resolution clearly approaches the theoretical. The central peak can be moved to the end of the screw and, in principle, double the resolution. However, the beam divergence would begin to limit the resolution, and it would become necessary to stop down the entrance or exit hole at all but the lowest wavenumbers.

14-8 ASYMMETRIC MODE

When the absorption cell is rotated into the fixed arm of the interferometer, the interferogram becomes asymmetric. The Fourier transform thus includes phase shifts from which the index of refraction can be calculated. The index information provides a useful supplement to the absorption data. Through the dispersion relations, the index of refraction can provide information about the absorption in regions where the absorption cannot be directly measured. This idea has been particularly useful in measuring line strengths in situations where it is difficult to obtain either sufficient resolution or weak enough lines to employ the usual techniques.

Figure 14-5 shows the index-of-refraction spectrum for a portion of the pure rotational band of CO. In the case of CO, the electronic contribution to the index of refraction is much greater than the rotational contribution. Since only the rotational part is of interest in this study, an arbitrary constant has been subtracted from the observed indices. In the figure, only changes in the index are meaningful.

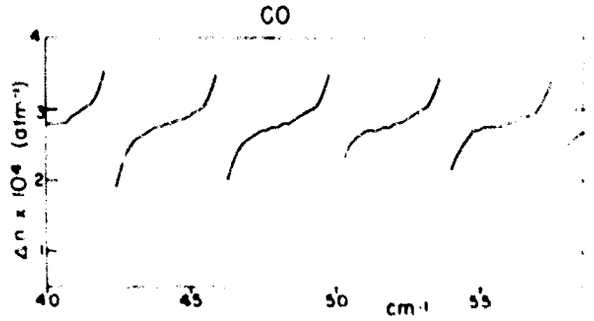


Figure 14-5. Index of Refraction of CO Normalized to One Atmosphere Pressure. An arbitrary constant has been subtracted from the values before plotting

The line strengths of CO have been calculated from this run by techniques described elsewhere.[2] They are in agreement with those calculated, assuming a dipole moment of 0.112 D, with about 10 percent uncertainty.[14] With only slightly more refined measurements, the stretching contributions should become observable, thus permitting a clear determination of the sign of the effective charge.

14-9 CARBON MONOXIDE LINE WIDTHS

As a further indication of the performance of the interferometer in applications to quantitative spectroscopy, the self-broadened half widths in the rotational spectrum of CO were measured. The CO lines are narrow and have half widths on the order of the resolution of the instrument. Ability to measure the widths of these lines depended strongly on the quality of the apparatus function of the interferometer.

There have been many studies of the CO line widths in the vibration-rotation band.[15,16] Theoretical calculations based on the Anderson theory have been carried out by Benedict and Herman.[17] Recently, Dowling measured the pure rotational widths using a lamellar grating interferometer.[18] In this latter work, effective half widths were measured directly from the spectrum. Theoretically calculated corrections were applied to remove the effect of the apparatus function and to correct for other sources of instrumental error. The Lorentz half widths were then calculated from the effective half widths.

Essentially the same procedures were followed in our analysis. The validity of these procedures depends on the assumption that any significant errors in the interferometer are sufficiently well understood that their effects on the line breadth can be calculated.

The rotational line widths were obtained from 13 runs made at pressures ranging from 490 to 745 Torr. The length of the cell was 22.5 cm and the temperature was maintained at 302±1° K. Two different filter combinations, as shown in Table 14-1, were used to cover the ranges 25 to 70 cm^{-1} and 50 to 130 cm^{-1}.

To optimize the signal-to-noise ratio, it is desirable to use the lowest resolution consistent with keeping the apparatus function corrections small. Using the formula given by Dowling,[18] the effect of an unapodized apparatus function on the width of a Lorentz line can be calculated. Using this formula, a maximum optical path difference of 10.4 cm, corresponding to a resolution of 0.096 cm^{-1}, was chosen. For this resolution, apparatus function correction was necessary only for the weakest lines observed.

Rotational transitions from $J=5$ (23.065 cm^{-1}) to $J=27$ (107.124 cm^{-1}) were observed. However, only the transitions from $J=7$ to $J=20$ were used. Below $J=7$, the signal-to-noise ratio was too low; above $J=20$, the lines were too weak.

Background runs were made by running empty cell interferograms at the lowest resolution consistent with reproducing all the remaining channel spectra. These background interferograms were extended to the same length as the sample interferograms and transformed. The transmittance curve defined by taking the ratio of the sample spectrum to the background spectrum has a straight base line without appreciable loss of signal-to-noise. The line profiles were then obtained by interpolating points at intervals of 0.004 cm^{-1} between the calculated points. The same program formed averages over suitably chosen intervals in the line wings to define the 100 percent transmittance level and calculate the width of the line at one half maximum absorption (the effective width).

Assuming the Lorentz shape, the transmittance in the neighborhood of an isolated absorption line is given by the expression

$$T(\sigma)=\exp[-k(\sigma)l], \tag{14-3}$$

where l is the absorption path length, and

$$k(\sigma)=\frac{S}{\pi}\frac{\gamma}{(\sigma-\sigma_0)^2+\gamma^2} \tag{14-4}$$

is the Lorentz shape; $S=S^0P$ is the line strength; σ_0 is the wavenumber of the line center; and $\gamma=\gamma^0P$ is the Lorentz half width. The effective half width γ_e is defined by

$$T(\sigma_0\pm\gamma_e)=\tfrac{1}{2}[1+T(\sigma_0)]. \tag{14-5}$$

This definition leads to an equation relating the effective half width to the Lorentz half width:

$$1+\left(\frac{\gamma_e}{\gamma}\right)^2=\frac{2x}{\ln\left[\frac{2}{1+e^{-2x}}\right]}, \tag{14-6}$$

where $x=N/2\pi\gamma$. Since x is independent of pressure, the effective half width is proportional to pressure, just as the Lorentz half width is. Each observed effective line width was corrected for the apparatus function, if necessary, and normalized to 1 atmosphere. Equation (14-6) was then solved for γ^0 by an iteration procedure. In this calculation, line strengths corresponding to a dipole moment of 0.112 D were used.[14] The half widths resulting from this procedure showed no systematic pressure dependence except for a few lines at $J>20$. For this reason, these lines were omitted from the analysis.

No other instrumental corrections were applied. Other calculable sources of error, such as the linear drift and the beam divergence, were analyzed and found to have a negligible effect on our widths. Many of these effects are much smaller for the Michelson interferometer than they are for the lamellar grating instrument.

Values of the Lorentz half widths for the lines $J=7$ through $J=20$ are listed in Table 14-2 and are plotted in Figure 14-6. The error bars are standard deviations and do not reflect any estimate of residual systematic errors. An uncertainty of 2 percent in locating the 100 percent base line leads to an error in the half width of about 3.5 percent or 0.0018 cm^{-1}/atm. This is consistent with the calculated standard deviations. The signal-to-noise ratio in the spectrum was about 50:1 above 35 cm^{-1}. At lower wavenumbers, it was somewhat poorer.

Table 14-2. Half Widths, γ^0, for the Pure Rotational Transitions of CO at $T=302°$ K

J	Wavenumber (cm^{-1})	γ^0 (cm^{-1} atm^{-1})	rms Deviation (cm^{-1})	Number of Lines Measured
7	30.748	0.0618	0.0015	3
8	34.588	0.0597	0.0014	6
9	38.426	0.0581	0.0019	6
10	42.263	0.0574	0.0018	5
11	46.098	0.0552	0.0011	6
12	49.932	0.0546	0.0010	6
13	53.763	0.0522	0.0020	7
14	57.593	0.0524	0.0015	8
15	61.420	0.0512	0.0015	11
16	65.245	0.0500	0.0016	10
17	69.068	0.0510	0.0013	5
18	72.888	0.0510	0.0013	5
19	76.705	0.0527	0.0015	4
20	80.519	0.0514	0.0020	4

Comparison with the results of Dowling[18] is good only for the lowest frequency lines. No evidence was seen of the discontinuity observed by Dowling at $J=9$. At high J values, our widths are greater than those of Dowling and show some tendency to increase with increasing J. The near infrared data show considerable scatter from one observer to another.[16] At

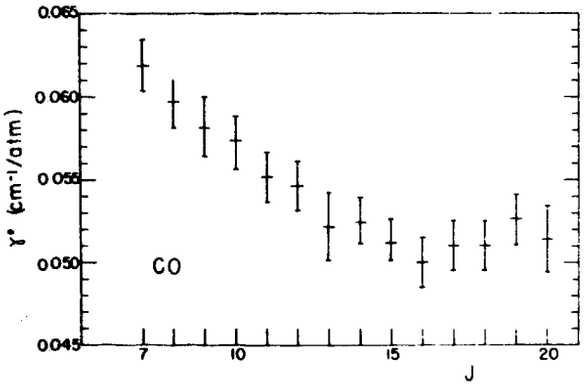


Figure 14-6. Lorentz Half Widths of the Pure Rotational Lines of CO at $T = 302°$ K. J is the lower state quantum number

least for the low J lines, we show reasonably good agreement with the measurements of Benedict, Herman, Moore, and Silverman[15] and with the theoretical calculations of Benedict and Herman,[17] assuming a quadrupole moment of 4×10^{-26} esu-cm^2.

The tendency of the line widths to increase at high J values indicates the presence of uncorrected systematic errors. It is unlikely, on theoretical grounds, that the lines actually become broader in this region. This behavior is probably due to small errors in the apparatus function which have not been taken into account in this analysis. The effective widths of these lines are not much greater than the width of the apparatus function and will be sensitive to small errors in the apparatus function. The low J lines are considerably broader, and appear to be more accurate.

Improvement in the high J lines will require two improvements in the analysis. The J dependence of the dipole moment matrix elements should be taken into account. Assuming that the effective charge is negative,[15] the strength of the $J = 20$ transition must be reduced by 8 percent and the width must be increased by 4 percent. As pointed out earlier, the asymmetric measurements should permit unambiguous determination of the direction of this correction. Secondly, the apparatus function should be measured directly, such as by observing very narrow lines. Realistically, one should not expect to be able to calculate an accurate apparatus function from the considerations used here.

14-10 CONCLUSION

In general, the interferometer has performed in accordance with expectations. It works very well at wavenumbers below 130 cm^{-1}. At higher wavenumbers the performance deteriorates. Spectra have been obtained to 240 cm^{-1} with good resolution but with noticeable deterioration in the apparatus function and in the signal-to-noise ratio. Satisfactory operation at high frequencies will depend on the development of an active control system for the optical path difference. Such a system is currently under construction. When this system is in use, the beam from a He-Ne laser will run through the interferometer in parallel with the infrared beam using the outer edge of the cat's-eyes. The stepping motor will advance the comparator screw a distance which is approximately a multiple of the laser wavelength. A servomechanism driven by the laser fringes will adjust the fixed cat's-eye to make the path difference an exact multiple of the laser wavelength.

The authors wish to acknowledge the continuing interest and assistance in the construction and operation of the interferometer of E. E. Bell, F. P. Dickey, and J. T. White.

References

1. Vanasse, G. A., and Sakai, H. (1967) *Progress in Optics VI*, North-Holland, Amsterdam, p. 261.

2. Sanderson, R. B. (1967) *Appl. Opt.* **6**: 1527.

3. Robinette, W. H., and Sanderson, R. B. (1969) *Appl. Opt.* **8**: 711.

4. Chamberlain, J., Gibbs, J. E., and Gebbie, H. A. (1969) *Infrared Phys.* **9**: 185.

5. Chamberlain, J. (1967) *J.Q.S.R.T.* **7**: 151.

6. Cuisenier, M., and Pinard, J. (1967) *J. Phys.* (*Suppl.*) **28**: 97.

7. Beer, R., and Marjaniemi, D. (1966) *Appl. Opt.* **5**: 1191.

8. Bell, E. E. (1967) *J. Phys.* (*Suppl.*) **28**: 18.

9. Rao, K. N., Humphreys, C. J., and Rank, D. H. (1966) *Wavelength Standards in the Infrared*, Academic Press, New York.

10. Filler, A. S. (1964) *J. Opt. Soc. Am.* **54**: 762.

11. Rao, K. N., and Williamson, J., *J. Mol. Spectry.* (to be published).

12. Rank, D. H., Eastman, D. P., Rao, B. S., and Wiggins, T. A. (1962) *J. Opt. Soc. Am.* **52**: 1.

13. Webb, D. U., and Rao, K. N. (1968) *J. Mol. Spectry.* **28**: 121.

14. Burrus, C. A. (1958) *J. Chem. Phys.* **28**: 427.

15. Benedict, W. S., Herman, R., Moore, G. E., and Silverman, S. (1962) *Astrophys. J.* **135**: 277.

16. Hunt, R. H., Toth, R. A., and Plyler, E. K. (1968) *J. Chem. Phys.* **49**: 3909.

17. Benedict, W. S., and Herman, R. (1963) *J.Q.S.R.T.* **3**: 265.

18. Dowling, J. M. (1969) *J.Q.S.R.T.* **9**: 1613.

Contents

15. A Computer Aided Design for an Interferometer Mirror Drive System

Joseph E. Hoffman, Jr.
Air Force Cambridge Research Laboratories
Bedford, Massachusetts

Abstract

An analog and a hybrid computer were used to simulate the performance of the mirror drive system of a Michelson interferometer with a long path difference. Separate simulations were performed for the stepping and the servo-controlled position-holding modes of operation. These simulations were performed to evaluate the system under a wide variety of conditions and to optimize the design.

15-1 INTRODUCTION

A Michelson interferometer with a 2-meter path difference is being developed. The mirror drive system operates in the step and integrate mode, in which the mirror is stationary during the measurement and then rapidly steps to the next position. The length of the step can be adjusted from 8 to 1000 fringes of the 6328 Å reference line. The stepping time must be minimized since it is an essentially wasted portion of the total measurement time.

The long path difference and fast stepping requirements led to the development of a two-part drive system*. The first part consists of a motor similar to the driver in a loudspeaker. This steps the mirror from one position to the next. The second part consists of an assembly which moves the motor mount at constant speed. These motions are combined and produce the modified staircase shown in Figure 15-1. A position-holding servo is engaged at the end of the

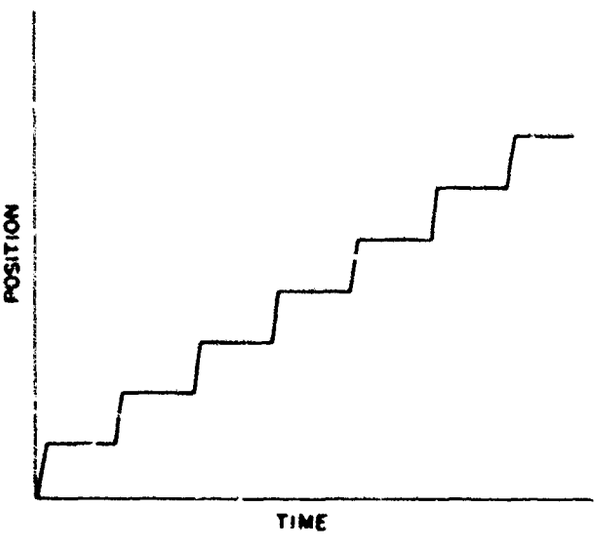


Figure 15-1. Mirror Position vs Time

* Pritchard, J. Private communication.

step and holds the mirror in the proper position during the measurement of the data.

Computer simulations were performed to obtain design and performance data and to optimize the system parameters. The stepping times and optimum braking points for several values of stepping distance, damping, and driving forces were found with an analog computer. A hybrid computer simulated the system in the position holding mode, and an automatic parameter search was made in order to find an optimum set of design parameters. These produce performance characterized by fast response, good stability, little ripple, and minimum error.

15-2 CHARACTERISTICS OF THE MIRROR DRIVE SYSTEM

The magnetic motor, power amplifier, feedback circuit, and mechanical load are shown in Figure 15-2.

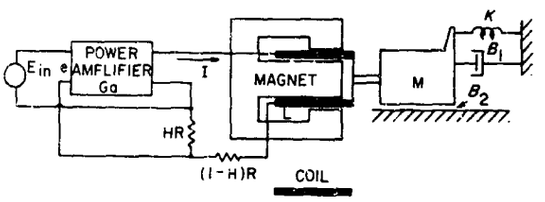


Figure 15-2. Magnetic Motor and Load

Table 15-1. Symbols Used in Magnetic Motor and Load Analysis

M	= Mass of the movable mirror and its carriage
B_1	= Viscous friction coefficient
B_2	= A nonlinear factor representing stiction and Coulomb friction
K	= Mechanical compliance
E_{IN}	= Source voltage
G_a	= Power amplifier gain
H	= Feedback factor
R	= Resistance of the coil circuit
L	= Inductance of the coil circuit
i	= Coil current
I	= Coil current (machine variable)
D	= Electromagnetic coupling coefficient
x	= Mirror displacement
X	= Mirror displacement (machine variable)
$\dot{x}$	= Mirror velocity
$\dot{X}$	= Mirror velocity (machine variable)
β	= Analog computer time scale factor
θ	= Tachometer constant
K_1	= Gain of the amplifier in the tachometer feedback loop
K_2	= Gain of the amplifier in the position feedback loop
$F(x)$	= Voltage versus position characteristic of the position reference

The symbols used in the analysis are defined in Table 15-1.

The motor exerts a constant accelerating force on the mirror until it reaches a predetermined position. Then a braking force is applied which brings the mirror to a stop at the end of the step. Several motors, capable of developing forces of up to 20 lb or 88.96 N, were designed and simulated. The values of R, L, i, and D were obtained from these designs.

The simulations include the effect of viscous friction, which consists of the damping produced by the magnetic flux cutting the metal coil form, and the effect of fluid damping. Stiction and Coulomb friction are also simulated. Provision was made in the program to include the effects of mechanical compliance, but it has been neglected in the initial sets of data.

15-3 SIMULATION OF THE MIRROR DRIVE SYSTEM IN THE STEPPING MODE

15-3.1 Derivation of the Computer Equations

The equation for the electrical portion of the system shown in Figure 15-2 is

$$G_a(E_{in}-HRi)=L\frac{di}{dt}+Ri+D\frac{dx}{dt}, \qquad (15\text{-}1)$$

or, when solved in terms of di/dt,

$$\frac{di}{dt}=\dot{i}=\frac{G_aE_{in}}{L}-\frac{G_aHRi}{L}-\frac{Ri}{L}-\frac{D}{L}\dot{x}. \qquad (15\text{-}2)$$

The equation for the mechanical portion of the system is

$$\frac{d^2x}{dt^2}=\ddot{x}=\frac{Di}{M}-\frac{B_1\dot{x}}{M}-\frac{K}{M}x-\frac{B_2}{M}. \qquad (15\text{-}3)$$

Equations (15-2) and (15-3) were programmed for solution on an analog computer. The time and amplitude scale factors were obtained from the theoretical values of stepping distance, stepping time, and maximum velocity when equal acceleration and braking forces and no damping were used. If the equations are scaled for a stepping distance of 1000

fringes of the 6328 Å reference, and are written in terms of the machine variables, the results are:

$$\left[\frac{\dot{I}}{10}\right]=100\left(\frac{.1G_a}{\beta L}\right)\left[\frac{E_{in}}{100}\right]$$
$$-100\left(\frac{G_aHR}{100\beta L}\right)\left[\frac{I}{10}\right]-\left(\frac{D}{100\beta L}\right)\left[\frac{\dot{X}}{10^{-1}}\right]$$
$$-10\left(\frac{.1R}{\beta L}\right)\left[\frac{I}{10}\right], \qquad (15\text{-}4)$$

and

$$\left[\frac{\ddot{X}}{10^{-1}}\right]=\left(\frac{100D}{\beta M}\right)\left[\frac{I}{10}\right]-\left(\frac{B_1}{\beta M}\right)\left[\frac{\dot{X}}{10^{-1}}\right]-\left(\frac{B_2}{\beta M}\right)[1]$$
$$-\left(\frac{K}{100\beta M}\right)\left[\frac{X}{10^{-3}}\right]. \qquad (15\text{-}5)$$

The analog computer program for solving Eqs. (15-4) and (15-5), with zero initial conditions, is shown in Figure 15-3. The portion of the circuitry below the dotted line is used to generate B_2, which has the form shown in Figure 15-4. Stiction, B_s, is effective until the velocity reaches a small value, b, then Coulomb friction, B_c, becomes effective. The relative magnitude of stiction, Coulomb friction, and the velocity at the transition point are adjustable.

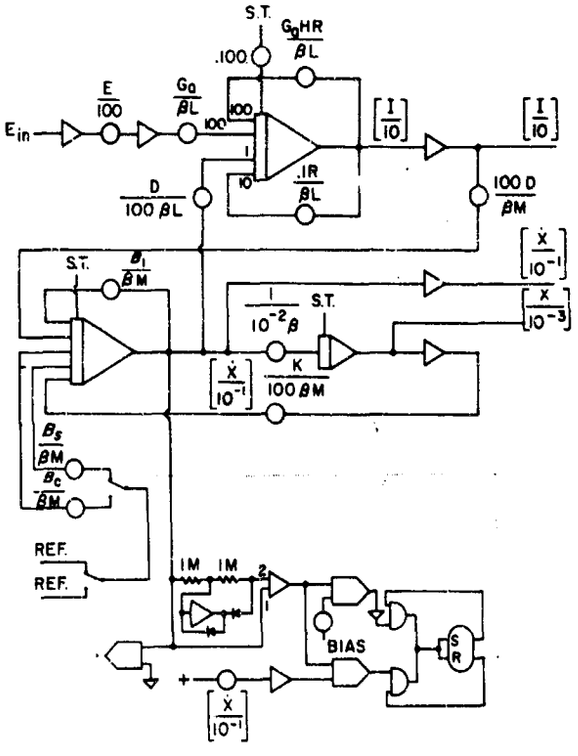


Figure 15-3. Analog Computer Diagram for Simulating the Stepping Mode of Operation

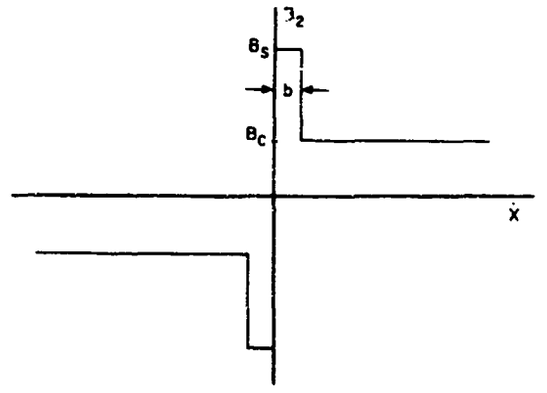


Figure 15-4. Function Representing Stiction and Coulomb Friction

15-3.2 Results of the Stepping Mode Simulation

An input voltage with a rectangular waveshape was used, and the response of the mirror was simulated for several values of motor time constant. The results are shown in Figure 15-5. The stepping distance was

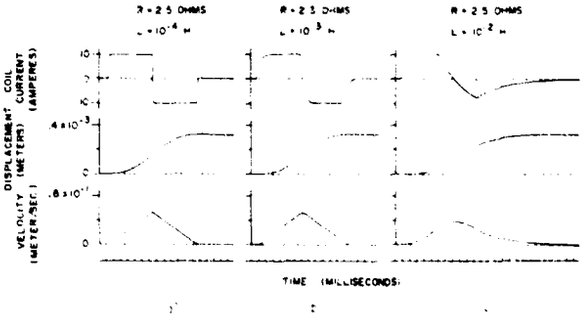


Figure 15-5. Response of the Mirror as a Function of R and L

1000 fringes of the reference line and no feedback around the power amplifier was used. Feedback effectively reduces the time constant, and the desired response can be obtained by a suitable choice of G_a and H. Feedback was used in the remainder of the simulations, and the values of G_a and H were chosen to give a response similar to the one shown in Figure 15-5a.

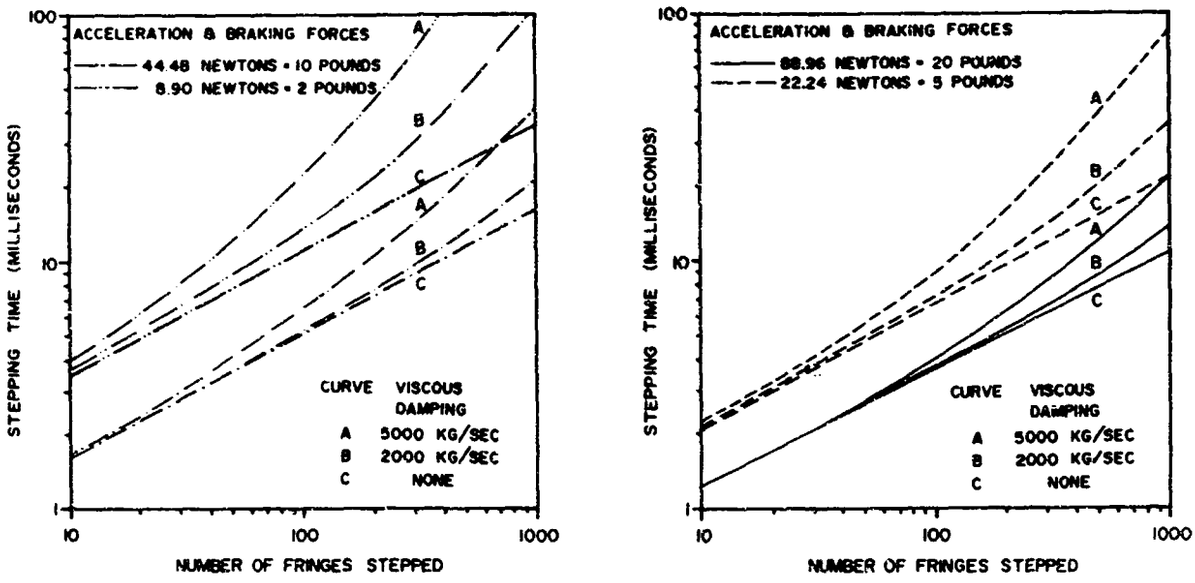


Figure 15-6. Stepping Time vs Stepping Distance

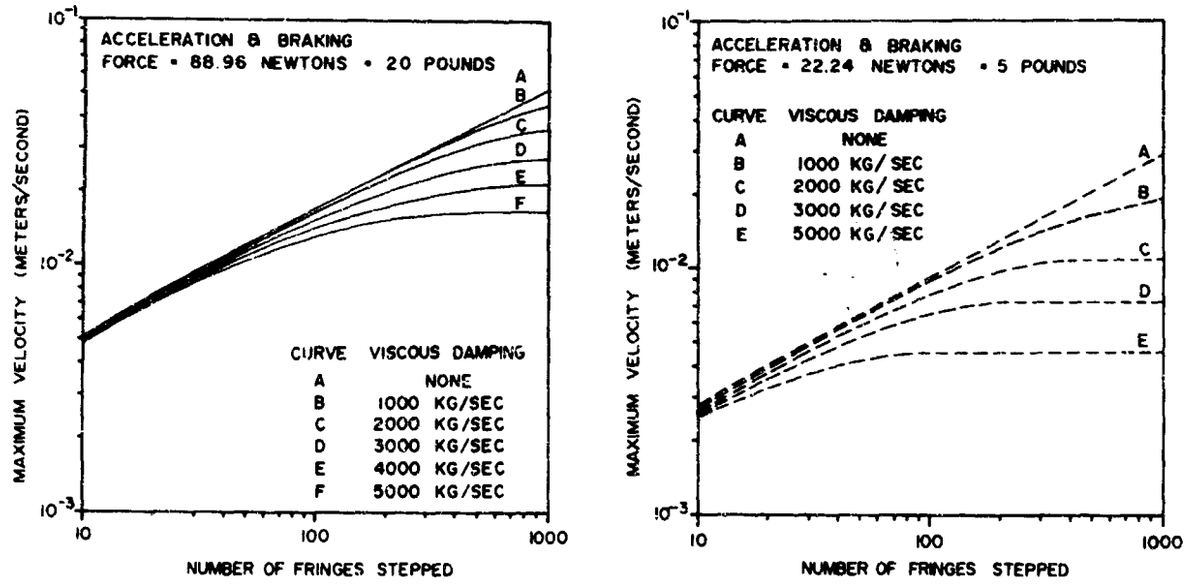


Figure 15-7. Mirror Velocity vs Stepping Distance

The fixed parameter values are: $G_a = 10$, $H = 0.9$, $D = 8.896$, $R = 2.5$ ohms, $L = 10^{-3}\,H$, and $M = 9.1$ kg. The variable parameters are: the stepping distance, viscous friction, stiction and Coulomb friction, and the acceleration and braking forces. The stepping times and maximum mirror velocities as functions of stepping distance are shown in Figures 15-6 and 15-7. The location of the braking point is critical if the mirror is to reach the end of the step and be caught by the servo. The optimum braking points are shown in Figure 15-8. Stiction and Coulomb friction effects have been neglected in these results.

When stiction and Coulomb friction are included, the response is slowed slightly as shown in Figure 15-9. This effect is not serious in the stepping mode because there are sufficient forces available to overcome stiction and the slight additional stepping time can be tolerated. However, in servo-controlled mode, stiction is an important problem because it produces position errors.

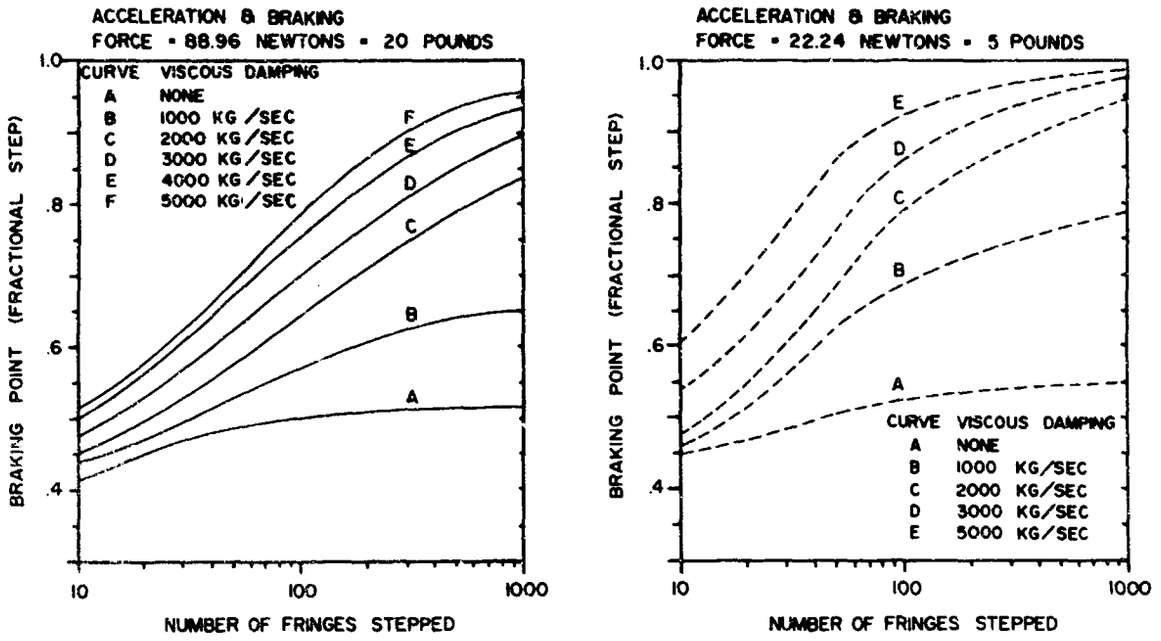


Figure 15-8. Braking Points vs Stepping Distance

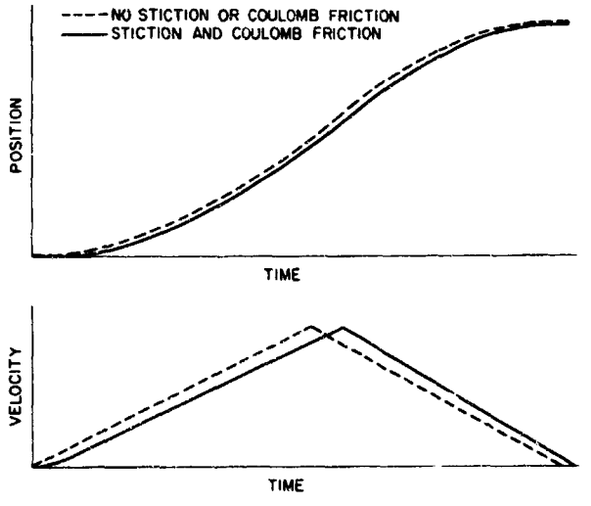


Figure 15-9. Effect of Stiction and Coulomb Friction in the Stepping Mode

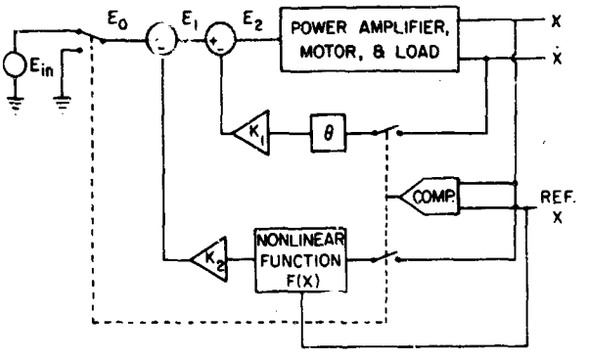


Figure 15-10. Block Diagram of the System in the Position-Holding Mode

15-4 SIMULATION OF THE MIRROR DRIVE SYSTEM IN THE POSITION HOLDING MODE

The position-holding servo is engaged when the mirror reaches the end of the step. A diagram of the modified system is shown in Figure 15-10. A comparator activates the servo by closing the tachometer and position feedback loops and by removing the driving source voltage. The reference for the position feedback is simulated by generating the function $F(x)$, shown in Figure 15-11, in the computer. The servo will operate properly as long as the position error is less than $|A|$. If the error exceeds $|A|$, the system will drive the mirror to the wrong point. Knowledge of the optimum braking point is important because late brake application leaves the mirror with too much speed and the servo cannot catch it before the position error exceeds $|A|$. Early brake application stops the mirror before it reaches the end of the desired stepping distance and the servo will not engage.

The equation for the electrical portion of the system is

$$G_a[-K_1\theta\dot{x}-K_2F(x)-HRi] = L\frac{di}{dt}+Ri+D\frac{dx}{dt}. \quad (15\text{-}6)$$

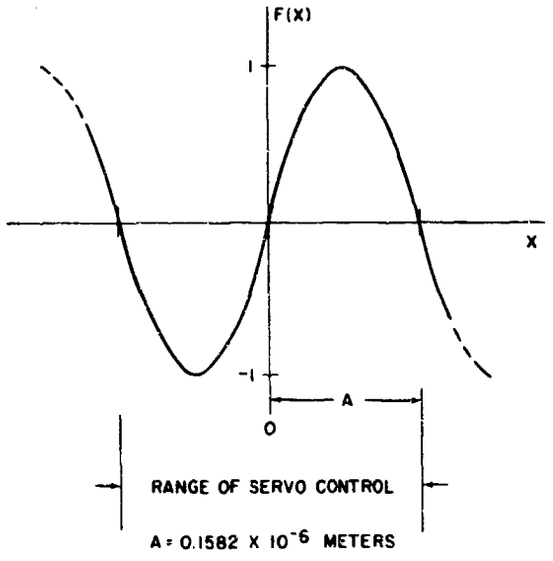


Figure 15-11. Position Reference

The addition of the feedback loops does not change the equation for the mechanical portion of the system and Eq. (15-2) still applies.

Several interesting problems may arise when this system is simulated. First, the mirror may have an initial velocity when the servo is engaged, which will not allow it to be caught. The allowable velocity is a function of the parameters of the system, which define a *velocity window*. The mirror can only be caught by the servo if its initial velocity lies within this *window*. Second, the time response and error are also functions of the parameters. The solution of these problems is difficult to obtain with a standard analog computer because of the vast number of parameters which must be investigated. A hybrid or a general purpose digital computer is more suited to problems of this type and can be programmed to find an optimum solution.

A hybrid computer was programmed to solve Eqs. (15-2) and (15-3). The diagram of the set-up is shown in Figure 15-12. Initial conditions used are

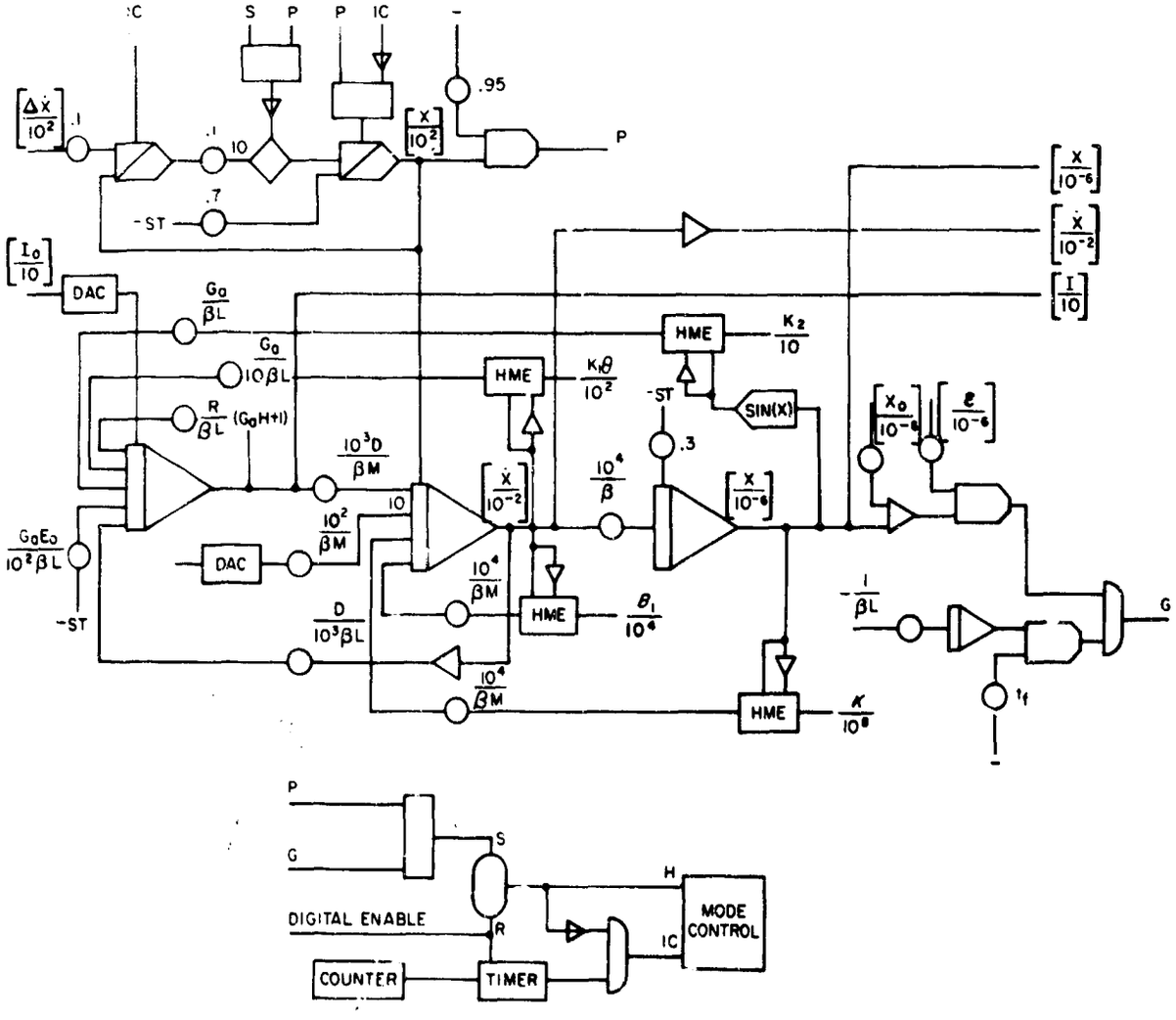


Figure 15-12. Hybrid Computer Diagram for Simulating the Position-Holding Mode of Operation

controlled by the digital part of the computer. The computer automatically varies the initial mirror velocity, the initial coil current, tachometer and position feedback gain, and the viscous damping. The parameters are varied according to a program, and a solution is obtained for each set of parameters. The computer automatically decides if the solution is valid. A valid solution is one in which the position error never exceeds $|A|$, the error converges to zero, and it reaches $|A|$ within a specified time. The computer records the parameters and indicates if the solution is valid. If so, the initial velocity is incremented and another solution is obtained. If an invalid solution occurs, the initial velocity is reset to its starting value and another parameter is changed according to the program.

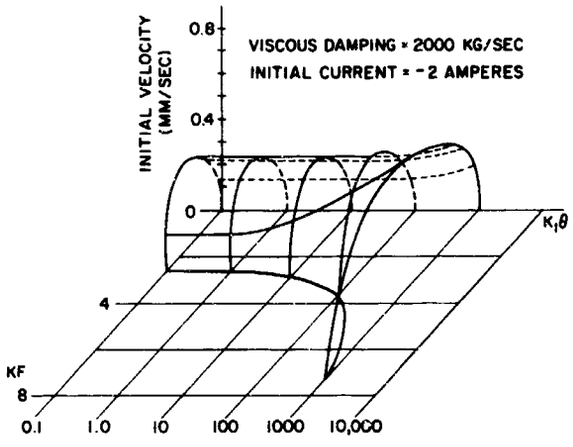


Figure 15-13. Example of a Velocity Window

An example of a *velocity window* is shown in Figure 15-13. The mirror can be caught if its initial velocity lies between the $KF-K_1\theta$ plane and the curved surface. Similar curves can be obtained for other parameter values.

Stiction and Coulomb friction cause position errors in the servo mode of operation. Several solutions of this problem are possible. The first involves a system to change the gain in the position feedback loop as a function of position error. This method will yield only partial correction. The second method involves correcting the path difference without further adjustment of the movable mirror. A piezoelectric crystal is used to mount a small mirror in the optical system of one of the retroreflectors. The mirror is servoed to adjust the path difference to the desired value.

15–5 SUMMARY

This paper illustrates a technique in which it is possible to use a computer to simulate an electromechanical mirror drive system and determine its performance characteristics. It is also possible to optimize the design by means of the simulation.

Discussion

Q. (J. V. Kline, Colorado School of Mines): Have you built this?

A. (J. Hoffman, Jr.): It is in the process; this system is presently under development. Computer simulation is useful to determine the optimum parameters for the mirror system and can be performed much easier than by the experimental method.

Contents

16. Modifications to the Aerospace Lamellar Grating Interferometer*

R. T. Hall, C. M. Randall, and J. M. Dowling
Space Physics Laboratory
The Aerospace Corporation
El Segundo, California

Abstract

Modifications to the lamellar grating interferometer are discussed including a tungsten source, an aluminum lamellar grating, and external adjustments for the final mirror before the exit aperture. Two planned modifications are the replacement of the Czerny-Turner mirrors by an off-axis paraboloid and the installation of a laser measuring system.

16–1 INTRODUCTION

Since the description[1,2] of the Aerospace Corporation lamellar grating interferometer was published in 1966, there have been several significant modifications to the instrument which have improved its performance. These modifications will be described and discussed. Finally, some major changes which are anticipated in the next several months will be mentioned.

16–2 DESCRIPTION OF CHANGES

First, the optical layout of the instrument will be reviewed. This is shown in Figure 16-1. The instrument is a double beam interferometer designed to operate in the far infrared beyond 250 cm^{-1}. The two beams are formed from the source and the two ellipsoidal mirrors, M_1 and M_2. One major change

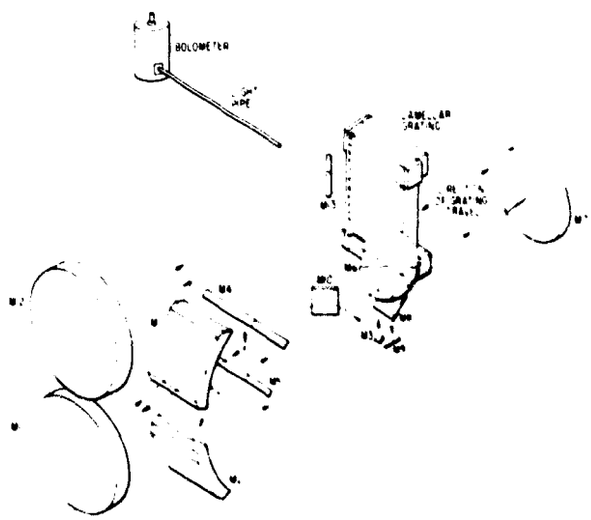


Figure 16-1. Optical Diagram of the Aerospace Corporation Lamellar Grating Interferometer

* This work was conducted under U.S. Air Force Space and Missile Systems Organization (SAMSO) Contract No. F04701-68-C-0066.

has been in this radiation source. The upper of the two beams is folded by mirrors M_4 and M_5 to pass through the sample area. The lower beam is a reference beam and passes unchanged under the sample area. It is folded up by mirror M_3. The two beams are then combined 180 deg out of phase by the rotating chopper wheel M_6. The ac component of this combined beam is the difference between the two beams. The combined beam is then refocussed by the spherical mirror M_7 and folded into the interferometer optics by mirrors M_8, M_9, and M_{10}. The entrance aperture lies immediately after mirror M_9. The beam is rendered parallel by the bottom Czerny-Turner mirror M_{11}, strikes the lamellar grating, and is then refocussed by the upper Czerny-Turner mirror M_{12}. Another of the major changes has been in the lamellar grating. An anticipated change is in the Czerny-Turner system, M_{11} and M_{12}. The refocussed beam is directed by mirror M_{13} to the exit aperture which is located on the end of the light pipe to the detector. The last major change has been to mirror M_{13}.

Perhaps the most significant change in the instrument has been the utilization of a tungsten coiled-coil lamp as a far-infrared source. This lamp is a commercial tungsten lamp* with half of the quartz envelope cut away. The remaining half is aluminized on the inside to eliminate channel spectra arising from reflections off the outside surface. The performance of this source is superior to a mercury arc above about 60 cm^{-1}. Spectra showing the intensity distributions of a mercury arc† and a tungsten lamp appear in Figure 16-2. The two beams of the instrument are collected from different portions of the source. Since the difference between the two beams is being considered, this has led to severe stability problems when using mercury arcs due to arc wander and envelope inhomogeneities. The use of a tungsten lamp has eliminated the stability problems of the mercury arc and has allowed spectra to be measured up to 250 cm^{-1}. Of course, for low-wavenumber investigations, the mercury arc must still be used. The stability problems are somewhat eased because the mercury arc can now be filtered at about 50 to 60 cm^{-1} so that only the plasma and not the hot envelope acts as a source.

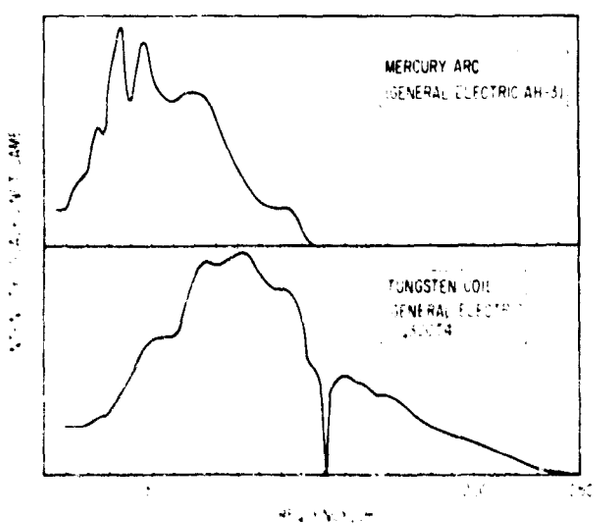


Figure 16-2. Far-infrared Intensity Distributions of a Mercury Arc and a Tungsten lamp

Since the bare tungsten coil will oxidize quite rapidly if operated in air, vacuum operation is a necessity. Since the instrument was already located in a vacuum chamber to eliminate water vapor interference, this proved no problem. The instrument is generally operated at better than 10^{-5} Torr vacuum. The operating characteristics of the lamp must be adjusted to account for the different heat load in the vacuum. Typically, the lamp is operated at about one third power (110 W) and a color temperature of 2500° K. Lifetimes of the order of 200 hr are usual. Failure of the lamps is preceded by a gradual change in cold resistance from 3.1 to 3.5 ohms.

Provision must be made to cool the sealed ends of the lamp. A water-cooled mount was designed for this purpose (see Figure 16-3). The mount and baffle

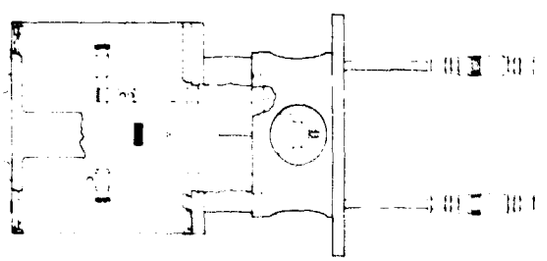


Figure 16-3. Source Mount and Housing

housing are designed to permit easy change between tungsten and mercury sources without having to realign the optics. This is done by having the source mount on a holder which accurately and reproducibly plugs into and is heat sunk to a water-cooled block. Electrical connections are made separately. The use of these plug-in mounts greatly reduces downtime when changing sources and eliminates the need for realignment at those times.

Another change made in the instrument is the replacement of the original Pyrex lamellar grating by an aluminum grating. This change was necessitated by a rather sudden deterioration of the optical figure of the Pyrex grating. In the past, objections have been raised against metal lamellar gratings

* General Electric 300 W Quartaline lamp, Q300T1H, 120V.
† General Electric H33A3 mercury lamp.

because of waveguide effects between the facets. However, investigations of the waveguide characteristics of the Pyrex grating revealed that this grating also exhibits a waveguide effect.[3] Because of the greater ease in fabrication and optical refiguring, the replacement grating was made from aluminum. The design of the new grating is shown in Figure 16-4.

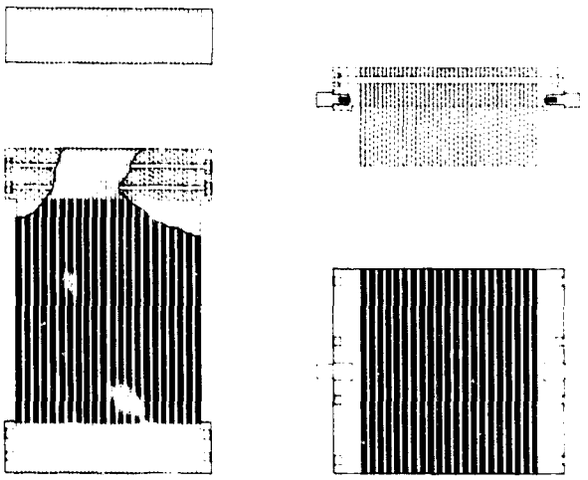


Figure 16-4. Design of the Aerospace Aluminum Lamellar Grating. The black areas are the optical surfaces

Figure 16-5. The Aerospace Aluminum Lamellar Grating in Place in the Instrument

The gratings are made from aluminum sheet stock, carefully machined and bolted together. The finished gratings were cold-stabilized prior to optical lapping. Stability tests were made on a small prototype version for a 45-day period.[4] These tests indicated adequate stability of the optical surfaces. The front surfaces of both halves of the grating were lapped separately and both have a slight cylindrical concavity. The mean facet-to-facet deviation from this cylindrical shape is 2 sec of arc, with the largest deviation of 8 sec of arc occurring at the outside facets. Since both halves are so nearly identical, the slight cylindrical shape presents no problems. The finished aluminum grating mounted in the instrument is shown in Figure 16-5.

One other modification which deserves mention is the provision for the external adjustment of the diagonal mirror M_{13} just before the exit aperture. This allows a more accurate centering of the zeroth order of the grating on the exit aperture. The adjustment is made by flexible cables attached to the three adjustment screws on this mirror. These cables can be turned from outside the vacuum tank by rotary motion feedthroughs mounted on the chamber wall. Prior to the installation of these adjustments, the centering had to be done in open air which is essentially opaque above 35 cm^{-1} due to water vapor absorption. Since the centering becomes more critical as the frequency is increased, misalignment for high frequencies (250 cm^{-1}) could occur if the adjustments were made at low frequencies (35 cm^{-1}). With the installation of these external adjustments, the final alignment can be done after the instrument is evacuated, thus eliminating the high frequency absorption.

16-3 FUTURE MODIFICATIONS

These are the principal modifications which have been made to the Aerospace interferometer. Two other modifications are planned for the near future. The first of these is the replacement of the Czerny-Turner mirrors, M_{11} and M_{12}, by an off-axis paraboloid. This will allow the instrument to be operated at near-normal incidence rather than at 10 deg off axis as is presently done with the Czerny-Turner system. This off-axis operation has been shown to broaden sharp, high frequency lines.[5] The installation of the off-axis paraboloid should decrease the instrumental broadening by a factor of 10 or more at 100 cm^{-1}.

The second modification planned is the installation of a laser for the measurement of optical path differences. The measurements are presently made by a

commercial linear inductosyn system which has an accuracy of $\pm 1.5\,\mu$m. Unfortunately, the system has a periodic error with an amplitude of $\pm 1.5\,\mu$m and a period of 2 mm. As in the ruling of diffraction gratings, periodic errors lead to "ghosts" in the spectrum. For example, a periodic error such as described above will give ghost lines at $\pm 5\,\mathrm{cm}^{-1}$ from a line at $50\,\mathrm{cm}^{-1}$ with an intensity of 2 percent of the parent $50\,\mathrm{cm}^{-1}$ line.[6] Spectra measured on the Aerospace interferometer have signal-to-noise ratios approaching this ratio of line-to-ghost intensities. Ghosts have been seen on a few spectra. A laser measuring system will eliminate these ghosts. Unfortunately, it appears that the "porch swing" fashion by which the moving grating is suspended has a great deal of vibration. This may necessitate a redesign of the drive and suspension system to allow a laser to be used.

These then are the main modifications of the Aerospace lamellar grating interferometer in a continuing series of modifications to improve its performance.

References

1. Hall, R. T., Vrabec, D., and Dowling, J. M. (1966) *Appl. Opt.* **5**: 1147.

2. Dowling, J. M., and Hall, R. T. (1967) *J. Phys., Colloque* C2, **28**: C2-156.

3. Dowling, J. M. (1969) *Linewidths in the Pure Rotational Spectrum of Carbon Monoxide*, Aerospace Rpt. No. TR-0066(9260-01)-4.

4. Dowling, J. M., Randall, C. M., and Mott, T. E. (1969) *Evaluation of a Lamellar Grating for the Aerospace Interferometer*, Aerospace Rpt. No. TR-0066(9260-01)-5.

5. Randall, C. M., and Dowling, J. M. (1969) *Evaluation of Parabolic Optics for the Aerospace Lamellar Grating Interferometer*, Aerospace Rpt. No. TR-0200(9260-01)-5.

6. Dowling, J. M. (1967) *Investigations in the Far Infrared With a Lamellar Grating Interferometer*, Aerospace Rpt. No. TR-100(9260-01)-7.

Discussion

Q. Is the plane of incidence and reflection at your lamellar grating parallel to the grooves, or not?

A. (R. T. Hall): Yes, the plane is parallel to the grooves of the lamellar grating. Since we are using a collimated beam, any other configuration would lead to shadowing of one set of facets by the other for any appreciable displacement.

Contents

17. New Instrumention for Far Infrared Fourier Spectrophotometers

L. W. Thorpe, D. J. Neale, and G. C. Hayward
Beckman R.I.I.C., Ltd
London, England

Abstract

Developments of the instrumentation available for use with a far infrared Michelson interferometer are discussed. Among these are a double-beam unit, a multipass gas cell, and a unit for obtaining reflection, transmission, and background interferograms without disturbing such instrumental conditions as vacuum and sample position.

17–1 INTRODUCTION

Fourier spectroscopy has developed over the last decade into a very powerful and well-used technique. Many of the early disadvantages have been eliminated or greatly reduced, such as the time delay and expense involved in the computation of the spectrum. The available instrumental techniques for sample handling at high or low temperatures have similarly been improved.

In general, for absorption studies the interferometer is run first on the background transmission of the instrument and then on the background transmission of the instrument and sample of interest. The transmission spectrum of the material is then obtained as the ratio of the transforms of the two interferograms.

One of the new pieces of instrumentation to be discussed in this chapter does away with the necessity of running two independent interferograms sequentially in time. This is the double-beam unit. A second piece of instrumentation is a tunable multipass gas cell with up to 13 m of path available in the cell. The third item is a module for obtaining both transmission and reflection spectra of a sample of interest as well as the background interferogram, without having to release the vacuum in the instrument.

17–2 THE REAL-TIME DOUBLE-BEAM SYSTEM

A real-time double-beam interferometer is a very convenient instrument. Among its advantages over a similar single-beam system it allows the background and sample interferograms to be obtained simultaneously and under identical instrumental conditions such as those of vacuum and detector sensitivity. The two interferograms that are obtained are sampled at identical positions of optical path difference, which produces a very convenient method of obtaining the complex refractive index of the sample in one run of the interferometer.

It is well known that a double-beam system gives more reproducible spectra since noise introduced by changes in source brightness and detector sensitivity

and by other instrumental effects affect both interferograms in exactly the same way. Provided the output interferograms have comparable dynamic ranges, the ratio of the two transforms will remove the effects produced by such fluctuations.

A double-beam system[1,2] developed at our Laboratories gives results that bear this out. The system uses a sum-and-difference technique whereby both beams are chopped at the same frequency but such that energy is received from each beam alternately, with complete obscuration between consecutive exposures. The resulting detector output waveform has one component at the chopping frequency proportional to the difference of the amplitude of the two input signals, and another component at twice the chopping frequency proportional to the sum of the amplitudes of the two signals. The sample and background interferograms can be separated from these two waveforms. Figure 17-1 (polypropylene

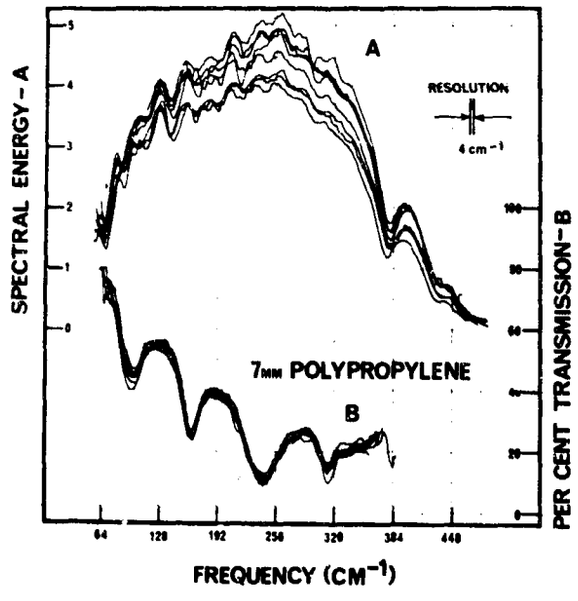


Figure 17-1. Comparison of Eight Ratioed Transforms for Polypropylene Sample in DB2 and Their Backgrounds

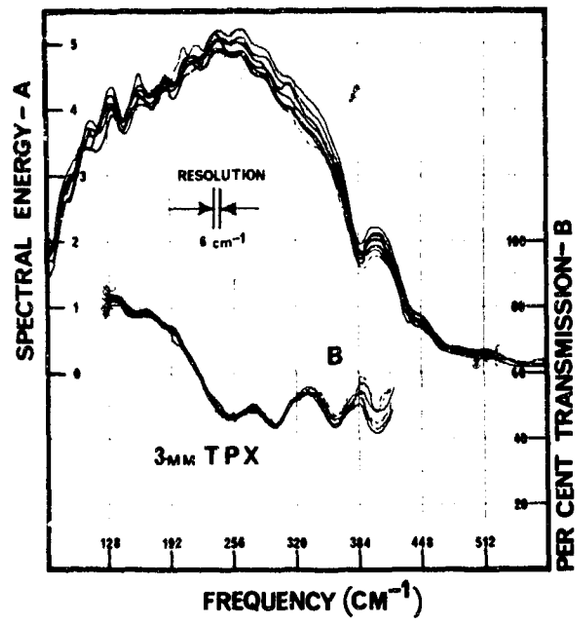


Figure 17-2. Comparison of Eight Ratioed Transforms for TPX Sample in DB2 and Their Backgrounds

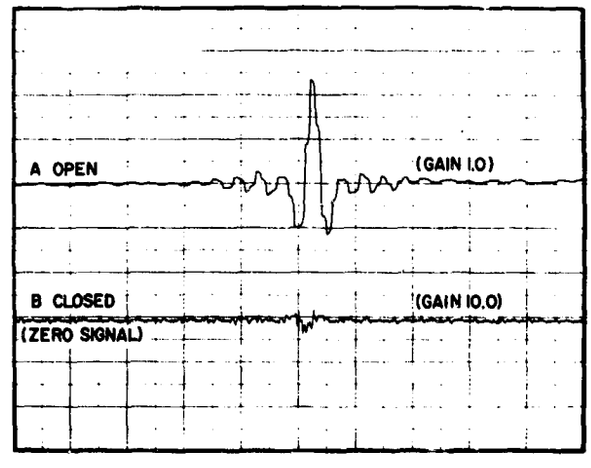


Figure 17-3. Cross-Talk Test (DB2)

sample) and Figure 17-2 (TPX sample) show the background transform and ratio of the sample and background transform for eight consecutive runs of the system. The improved reproducibility in the region of good signal levels is immediately obvious from a comparison of plots A and B in both cases.

The double-beam system suffers from a cross-talk problem, that is, when one channel is closed and the other open, energy from the open channel is observed in the output signal representing the closed channel, as shown in Figure 17-3. The measured cross-talk is found to be dependent on the amplitude of the second harmonic component obtained in the detector output when the radiation is coded with the normal square-wave type of chopping. Hence, the summing and differencing of the amplitude of the fundamental and second harmonic of the chopping frequency introduces false energy into both channels, with a resulting error in the measured spectrum. Only by careful selection of the detectors can the cross-talk be reduced to below 1 percent, which is considered adequate for normal working. A second double-beam system that has

since been developed should not present the cross-talk problem since it does not depend on a sum-and-difference technique to produce the two electrical outputs. Results will be published later.

With either double-beam system, it is possible to measure the complex refractive index of a parallel-sided sample[3] placed in one of the two beams of the interferometer. The following notation is used.

- $F(\nu)$ = energy function incident n sample as a function of wavenumber
- $T(\nu)$ transmitted energy function as a function of wavenumber
- $\mu(\nu)$ refractive index of sample as a function of wavenumber
- t thickness of parallel-sided sample (in centimeters)
- Δx sampling interval (in centimeters)
- E effective extra optical path length due to a parallel-sided sample of thickness $t=2t[\mu(\nu)-1]$ for all wavenumbers

For the background the interferogram obtained as a function of the sample point n is:

$$f(n\,\Delta x)=\int_{\nu_0}^{\nu_1}F(\nu)\cos(2\pi\nu n\,\Delta x+2\pi q\nu\,\Delta x)\,d\nu, \quad (17\text{-}1)$$

where $|q|\leq 0.5$ and represents the error in the sampling of the background zero path difference position $n=0$.

With the sample introduced into the Michelson interferometer, the interferogram obtained as a function of the sample point n is:

$$f_1(n\,\Delta x)=\int_{\nu_0}^{\nu_1}F(\nu)T(\nu)\cos\{2\pi\nu n\,\Delta x+2\pi q'\nu\,\Delta x+2t[\mu(\nu)-1]\}\,d\nu, \quad (17\text{-}2)$$

where

$$[\mu(\nu)-1]=\frac{\epsilon}{4\pi\nu t}-\frac{(n_2-n_1)}{2t}\Delta x-\frac{(q'-q)}{2t}\Delta x;$$

$|q'|\leq 0.5$ and represents the error in the sampling of the zero path difference region, and ϵ is phase difference.

Since both interferograms are obtained from the same interferometer at the same position of optical path difference and at the same instant of time, the two phase functions inside the cosine terms in Eqs. (17-1) and (17-2) are related and can be compared. The phase existing in the background interferogram at the position taken as the zero path difference sample point n_1 is:

$$\text{bkgd phase}=2\pi\nu(n_1\,\Delta x+q\,\Delta x); \quad (17\text{-}3)$$

the phase existing in the sample interferogram at the position taken as the zero path difference sample point n_2 is:

$$\text{sample phase}=2\pi\nu(n_2\,\Delta x+2t[\mu(\nu)-1]+q'\,\Delta x). \quad (17\text{-}4)$$

The phase difference as a function of frequency and zero path difference sample point numbers is:

$$\epsilon=2\pi\nu(n_2-n_1)\,\Delta x+2\pi\nu\cdot 2t[\mu(\nu)-1]+2\pi\nu(q'-q)\,\Delta x. \quad (17\text{-}5)$$

The computation of this experiment is done in two steps.

1. The computer recognizes each of the two positions n_2 and n_1 as the maximum element in its particular interferogram; it then finds the average value of the refractive index, given by:

$$\text{avg R.I.}=\frac{(n_2-n_1)}{2t}\Delta x. \quad (17\text{-}6)$$

2. During the phase correction of both interferograms the two phase functions are calculated, each based about its own zero path difference position. Each phase function has a component that represents the mis-sampling of the zero path differences position plus any phase changes with frequency due to misalignment of the interferometer. Since only the sample phase function contains information about the refractive index of the sample, subtraction of the two

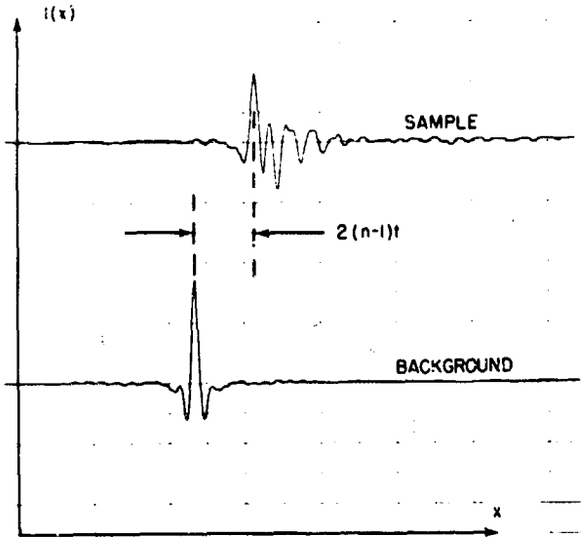


Figure 17-4. Asymmetric Interferogram of Teflon and the Corresponding Symmetric Interferogram of the Background

phase functions removes two causes of error from the refractive index measurements. Accurate location to a fraction of a sampling interval, of either or both of the zero path difference positions, therefore becomes unnecessary.

The real part of the complex refractive index is obtained from the ratio of the Fourier transforms of the phase-corrected sample and background interferograms; the imaginary part is obtained from the difference of the two phase functions and from other relevant factors described in the equations.

Figure 17-4 shows the symmetric background and asymmetric sample interferograms. Figure 17-5 shows the results after computation of eight such scans of the interferometers, with a 147-μm-thick sample of Teflon placed in the interferometer. In the transform of the sample and background interferograms the spread is seen to be about 10 percent; in the ratio of the two transforms, in the regions of good signal levels, the spread is about 2 percent. The changes in refractive index are seen to be associated with the absorption lines of Teflon, and the results agree with previous classical refractive index measurements.

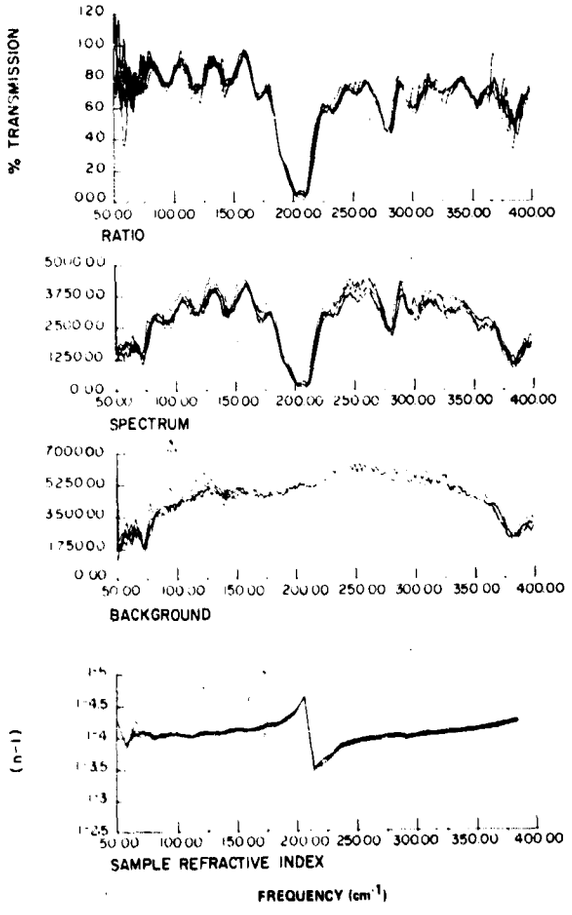


Figure 17-5. Eight Asymmetric Sample Interferograms (147-μm-thick Teflon) and Their Eight Symmetric Background Interferograms, Shown With Digital Computation and Sample Refractive Index

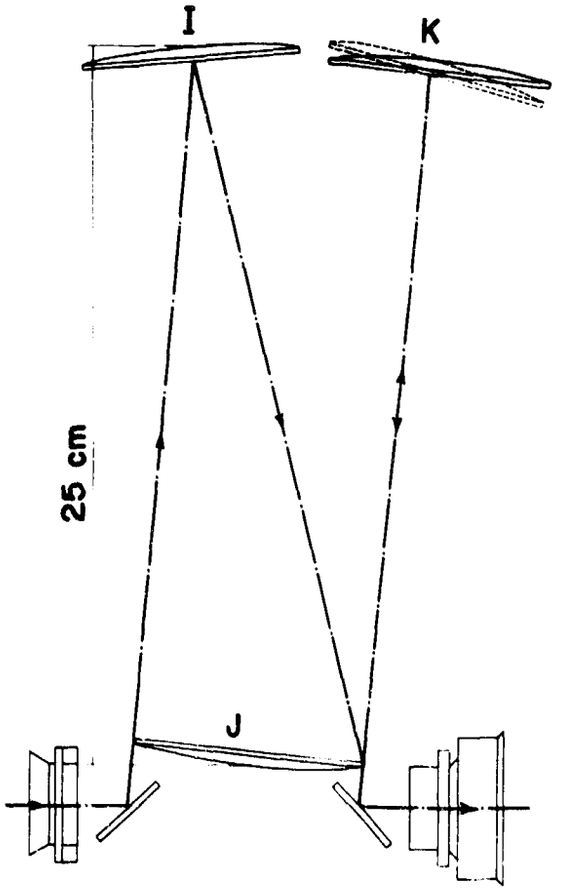


Figure 17-6. Optical Layout of Tunable Multipass Gas Cell for One Traverse of the System, Giving 1 m Transmission in the Gas

17-3 MULTIPASS GAS CELL

The multipass gas cell available as a modular extension to the existing Fourier spectrometer is of a type developed by White,[4] with an optical layout as shown in Figure 17-6. The optical components consist of three spherical mirrors, all having the same radius of curvature, and spaced apart by distances equal to this radius. The system can be tuned from 1 m to more than 10 m of transmission in the gas and the number of passes is determined by the relative positions of the centers of curvature of the mirrors K and I on the mirror J, as shown in Figure 17-7. The

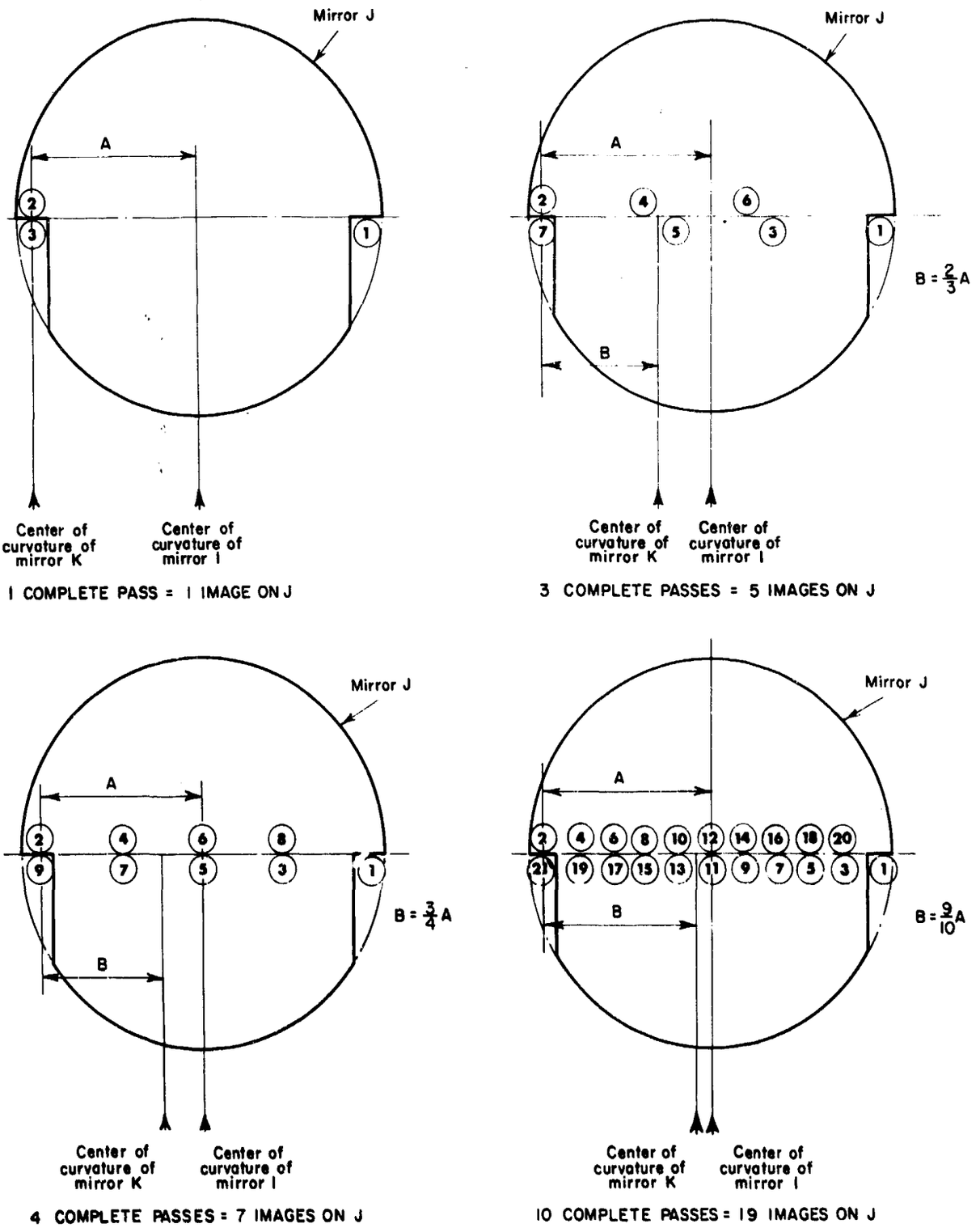


Figure 17-7. Mirror J, of Multipass Gas Cell, and Number and Position of the Images for 1, 3, 4, and 10 Passes Through the Cell

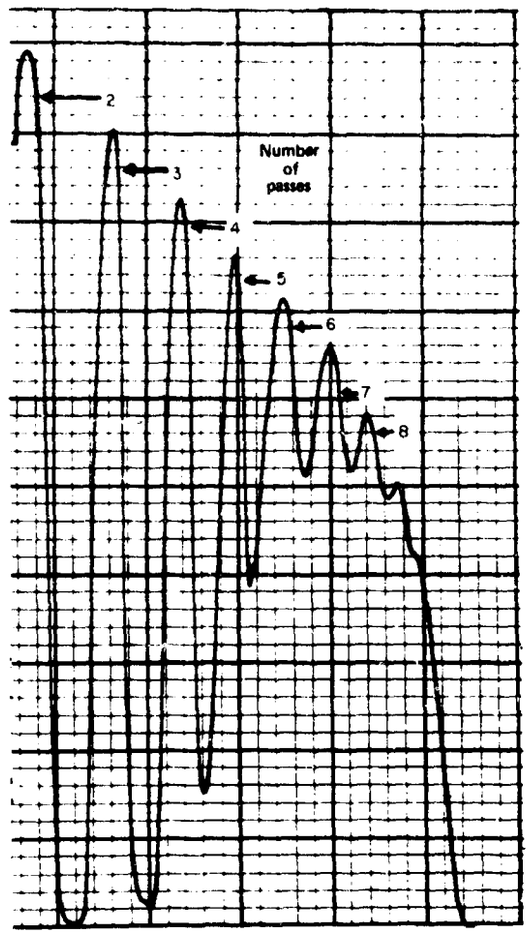


Figure 17-8. Multipass Gas Cell Throughput With 10-mm Source Aperture

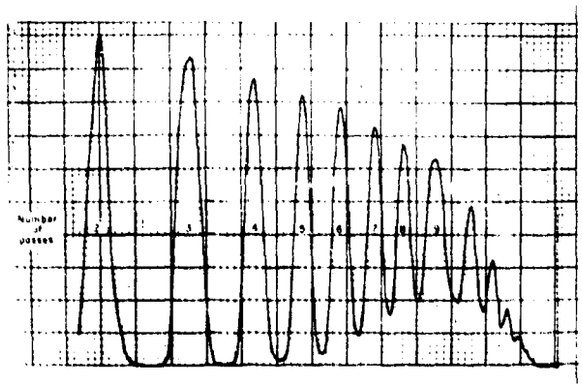


Figure 17-9. Multipass Gas Cell Throughput With 3-mm Source Aperture

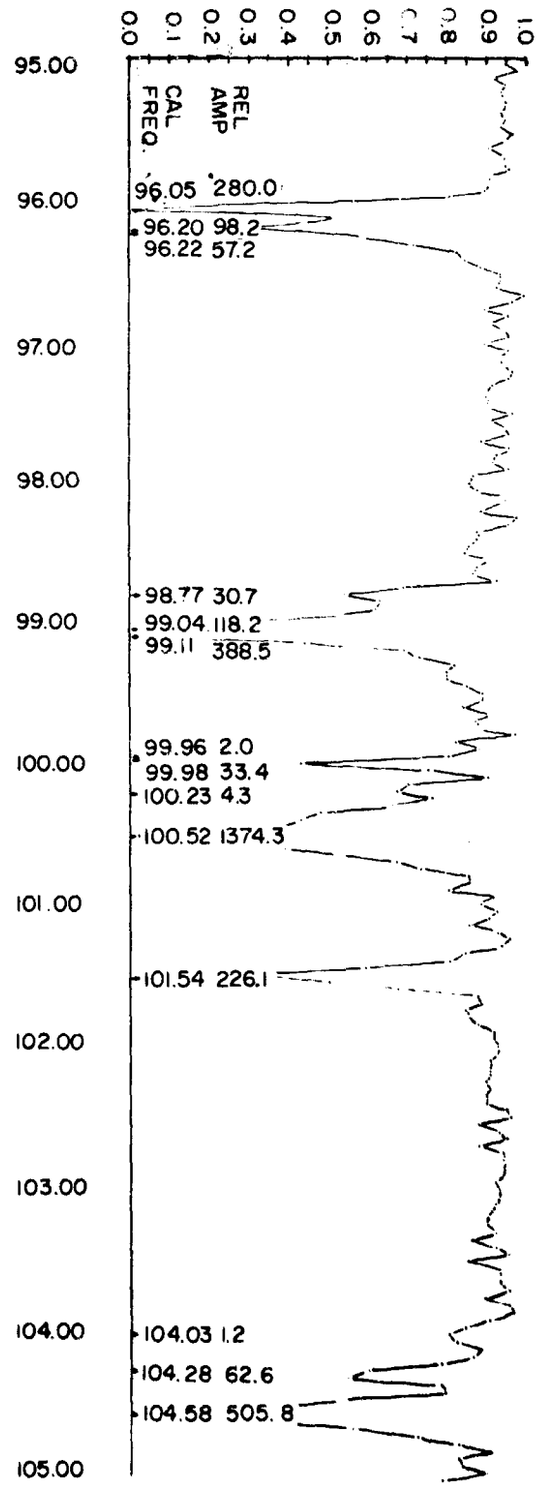


Figure 17-10. High-resolution Spectrum of H_2O Obtained by Using 1 m of Path in the Multipass Gas Cell and 9 cm of Optical Path Difference (Results were obtained by means of a point-by-point Fourier transform procedure.)

cell is designed to be used at pressures up to 2 atm and as high a vacuum as is available.

A maximum in the transmitted energy is obtained (see Figure 17-7) when $A/(A-B)=N$, an integer representing the number of passes of the cell. The limit of the number of passes occurs when $2(A-B)=S$, the diameter of the individual images on the mirror J, or,

$$N_{\text{limit}}=2\,A/S.$$

Figure 17-8 shows the number of passes obtained as the mirror K is adjusted, using an image diameter of 1 cm on the mirror J.

Figure 17-9 shows the results of a similar test with an image diameter of 3 mm on the mirror J.

Figure 17-10 shows a digitally computed high-resolution spectrum of water vapor obtained by using the gas cell and a standard Michelson interferometer whose drive is capable of a maximum of 18 cm of optical path difference. This spectrum shows the splitting of the 96 cm^{-1} and 99 cm^{-1} doublets to better than 0.1 cm^{-1}. Results were obtained by means of a point-by-point Fourier transform procedure.

Figure 17-11 shows an FTC-300 analog computer spectrum of CF_3Br at approximately $\frac{1}{5}$ atm and 9 passes in the cell. This gas is very volatile and was distilled five times to remove the water present in the gas as obtained from the manufacturer.

Figure 17-12 shows three FTC-300 comparison spectra of C_6F_6, for 2, 6, and 9 passes in the cell, and demonstrates the advantages in sensitivity to be obtained by the extended transmission available in the gas cell.

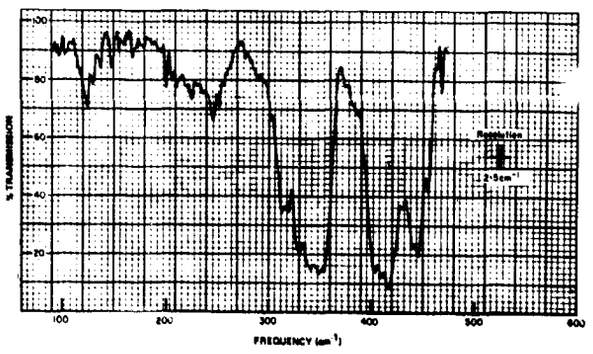


Figure 17-11. An FTC-300 Ratioed Spectrum of CF_3Br Obtained by Using 9 m of Transmission in the Multipass Gas Cell

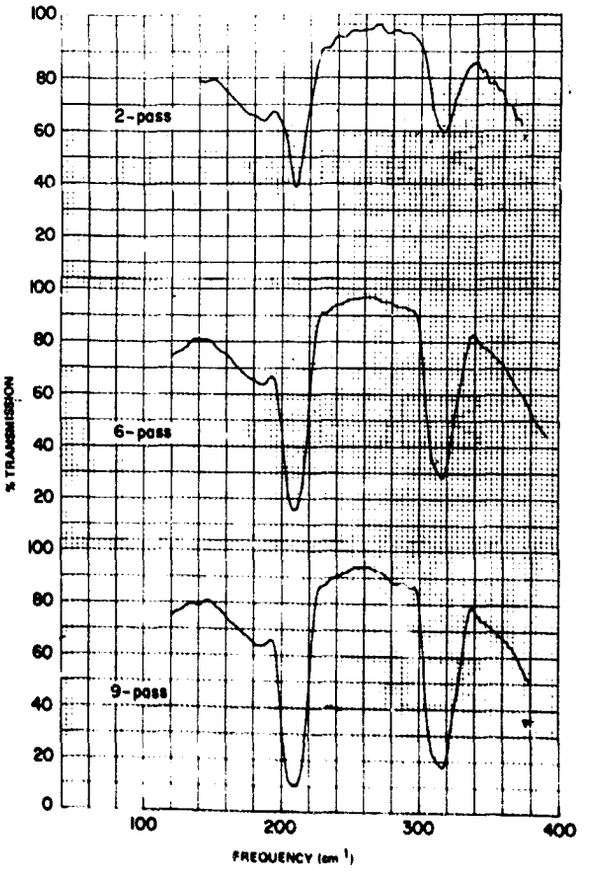


Figure 17-12. Three FTC-300 Spectral Plots of C_6F_6, for 2, 6, and 9 Passes in the Multipass Gas Cell

17-4 RANGE EXTENSION OF THE FS-720 FAR INFRARED MICHELSON INTERFEROMETER TO HIGHER FREQUENCIES

Their large energy throughput and high-resolution capability has aroused considerable interest in the feasibility of extending the range of far infrared Michelson interferometers to higher frequencies. Advantages to be gained are the possibility of overlapping the range covered by relatively inexpensive dispersive instruments and refractive index measurements with the interferometer used in the asymmetric mode. A further convenience is the modular construction of these instruments.

The wavelength range covered by the FS-720 Michelson interferometer is to a large extent determined by the spacing of the moire grating it uses for sensing the displacement of the moving mirror. The grating has a 4-μm period, and the interferogram is normally sampled once for every moire fringe, that is, at every 8-μm increment of optical path difference. Hence, according to the sampling theorem, for unambiguous transformation of the interferogram the spectral bandwidth of the instrument must be less than 625 cm^{-1}, and since the instrument uses only lowpass filters, the highest frequency that can be covered is theoretically 625 cm^{-1}.

The normal range of the interferometer is from 10 to 500 cm^{-1}. The instrument background for use at the higher frequencies is provided by interference within a 6-μm-thick Mylar beam splitter; all radiation of frequencies higher than approximately 500 cm^{-1} is removed by the black polyethylene field lens. It is possible to extend the frequency range to slightly higher values by using thinner beam splitters, and the one most commonly used is of 3.5-μm-thick Mylar. Unfortunately, since Mylar shows absorptions near 380, 440, and 510 cm^{-1}, the spectral background that it provides is rather unsatisfactory in this region. A better spectral background, which shows no absorption at these frequencies, is obtained by using polypropylene beam splitters.

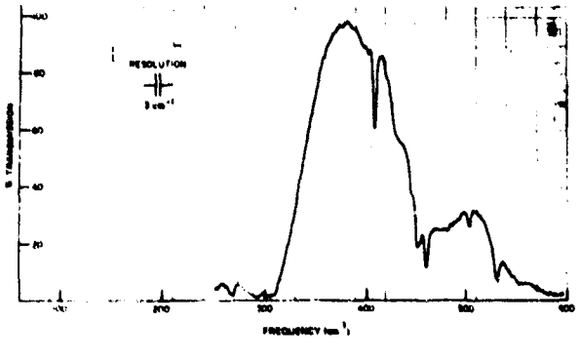


Figure 17-13. Single-beam Spectrum of an Antimony Complex (PH_2SbCl_2O, in Polyethylene on a Spectral Background Optimized for the High-Frequency End of the Instrument Range (SI, 8 μm; optical filter, KBr)

Figure 17-13 shows a single-beam spectrum of a polyethylene disk containing an antimony complex against a spectral background optimized for the high-frequency end of the instrument range. The spectrum was obtained with a polypropylene beam splitter 10 μm thick and a KBr filter to remove long-wavelength radiation. Figure 17-14 shows the ratioed transmission spectrum of the same disk as in Figure 17-13; the high-frequency limit of this spectrum is about the practical limit for 8-μm sampling when black polyethylene is used as a filter. In both Figures 17-13 and 17-14 the spectrum was plotted with the FTC-300 computer on the $\times 1\frac{1}{2}$ range, which displays frequencies up to 96 percent of the alias values.

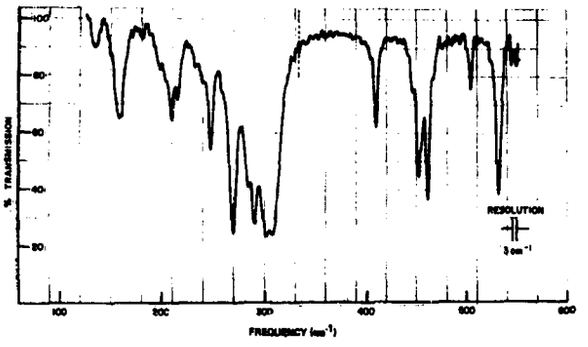


Figure 17-14. Ratioed Transmission Spectrum of a Polyethylene Disk Containing the Same Antimony Complex as in Figure 17-13.

To investigate the feasibility of extending the instrument range further, it was necessary to reduce the sampling interval. This was achieved by modifying the moire amplifier electronics so that the interferogram was sampled twice per moire fringe, that is, at 4-μm intervals of optical path difference. The accuracy of sampling was comparable to that of the standard moire system, with the alias frequency increased to 1250 cm^{-1}.

The interferometer lends itself readily to range extension since the optics are all reflecting and beam splitters are readily interchangeable. The beam splitters we tried were of three different materials, down to the following thicknesses: Mylar, 2.5 μm; polypropylene, 5 μm; and polyethylene, 12 μm. All these materials were strong enough to be used interchangeably, although thin polyethylene films tended to stretch slightly with use and could not be interchanged more than about six times.

The tilt adjustments that were provided for the fixed mirrors were found to be fine enough for satisfactory optical alignment at the higher frequencies. The black polyethylene field lens was removed and an 8-μm or 12-μm lowpass filter was inserted near the

sample position. Although the optics were thus not optimized, there was so much radiation from the source at higher frequencies that it was necessary to provide for additional attenuation (in the range 10 to 20 percent transmission) in order to avoid overloading the detector, even when the instrument was operated at the smallest (3 mm) source aperture.

The smaller radiative throughput of beam splitters of polyethylene and polypropylene, which have lower refractive indices than Mylar, is thus unimportant. It was possible to obtain satisfactory interferogram modulation depths for all beam splitters, and all films other than polyethylene. One specimen of polypropylene was sufficiently uniform to give good visible fringes. An effect reported by other workers,[5] that the interferogram fringe contrast was improved by removing the low-frequency content of the spectral background with alkali halide transmission filters, was also observed by us.

The results of measurements with various beam splitter and filter combinations are shown in Figures 17-15 through 17-18.

Interferograms for spectra at frequencies less than 500 cm^{-1} are normally scanned at 5 μm/sec with an electronic time constant of 0.5 sec, and it was necessary to lower the drive speed by a factor of 2 when operating at these high frequencies. The running

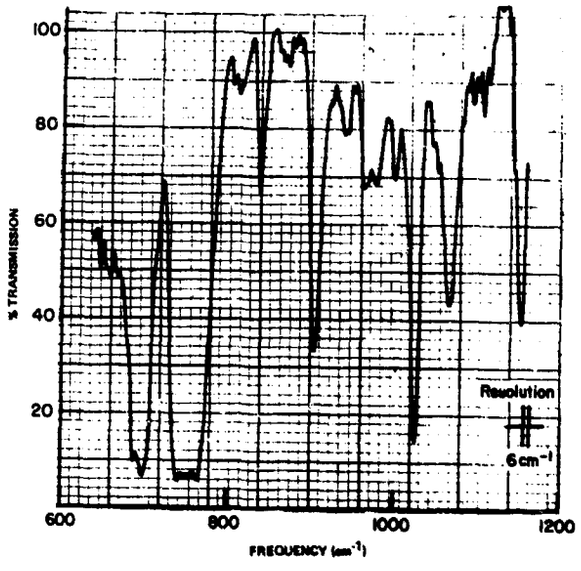


Figure 17-15. Ratioed Spectrum of a Polystyrene Film Using a 5-μm Polypropylene Beam Splitter (SI, 4 μm)

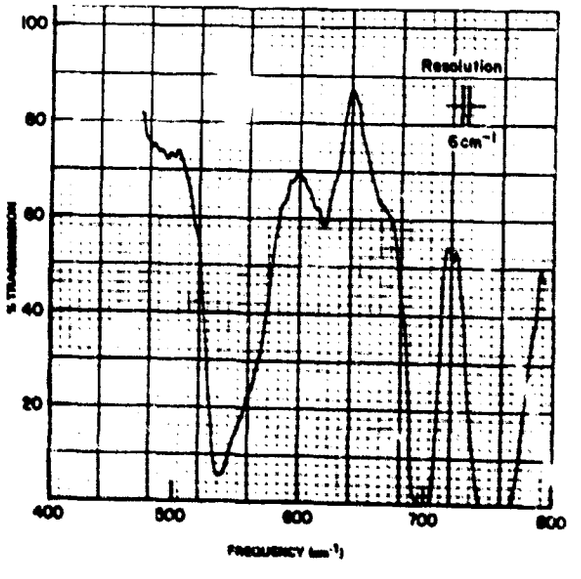


Figure 17-16. Ratioed Spectrum of a Polystyrene Film Using a 3.5-μm Mylar Beam Splitter (SI, 4 μm)

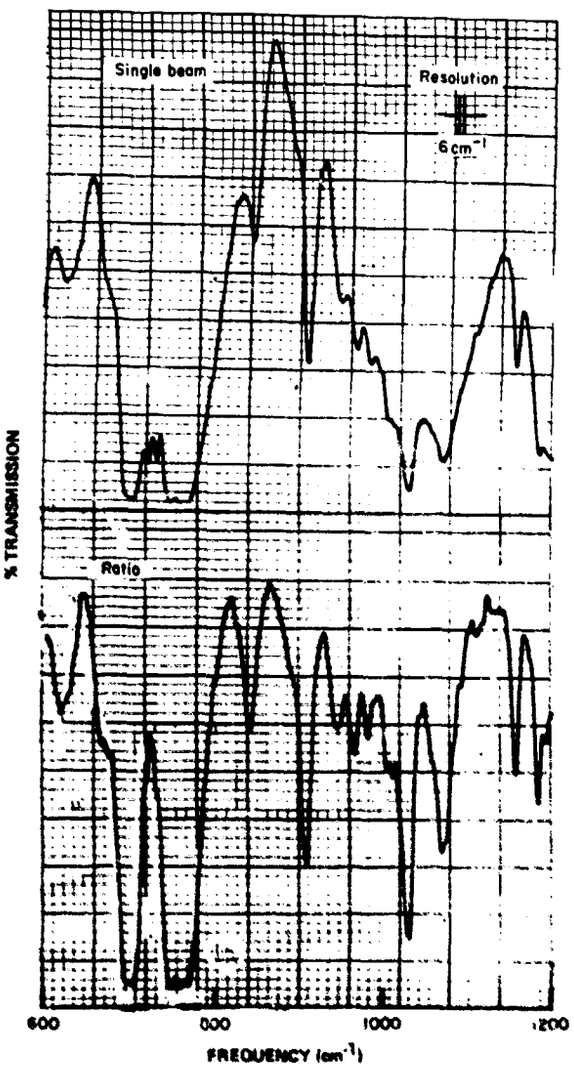


Figure 17-17. Ratioed Spectrum of a Polystyrene Film Using a 12.5-μm Polyethylene Beam Splitter (SI, 4 μm)

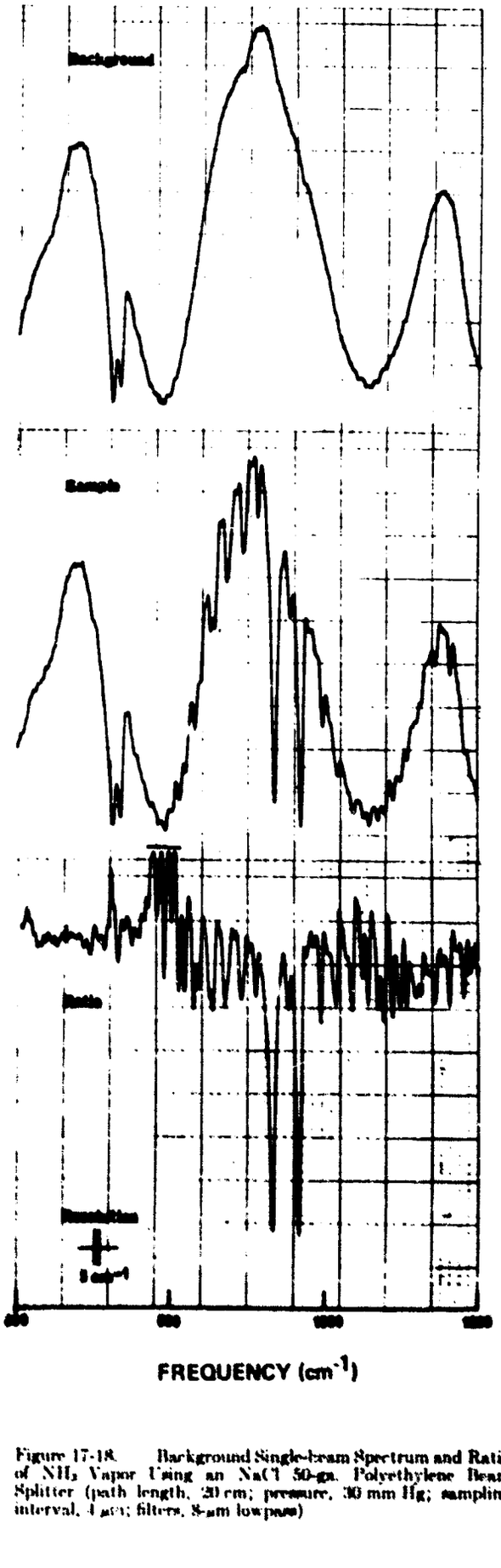


Figure 17-18. Background Single-beam Spectrum and Ratio of NH_3 Vapor Using an NaCl 50-gm. Polyethylene Beam Splitter (path length, 20 cm; pressure, 30 mm Hg; sampling interval, 4 μm; filters, 8-μm lowpass)

time of interferograms for spectral resolutions of 5 to 6 cm^{-1}, the highest that could be achieved with the limited storage capacity of the analog computer, was 20 min, this time being dictated by the response time of the Golay detector.

The results show that polypropylene is a satisfactory beam splitter for frequencies up to 800 cm^{-1} and virtually free from absorption features. This result is at variance with those of other workers who apparently find that absorption in the filters or lens materials is ascribable to the beam splitter. Mylar is less satisfactory for this range since besides absorption at the lower frequencies there is intense absorption near 720 cm^{-1}. Of the materials studied, polyethylene is the only one suitable for use in the region 800 to 1200 cm^{-1} (Figure 17-19), although the thickness of film used—which is about the thinnest that can be tolerated in practical use—is such that this region cannot be covered by a single beam splitter "hoop" (see Figure 17-20). Possibly because the polyethylene beam splitter was not uniform, the spectral intensity at the beam splitter minima was not zero, and so it was possible to ratio spectra with reasonable success in the region of these minima.

The results suggest that with only minor modifications the wavelength limit of the interferometer can be extended to at least 800 cm^{-1}, and indicate the splitter/filter combination necessary to optimize performance in this region.

17-5 REFLECTION-TRANSMISSION ATTACHMENT

This module is designed to fit onto the existing FS-720 interferometer such that the instrumental background, reflection, and transmission of the sample are available without having to disturb either the sample once it has been positioned, or the instrument once it has been evacuated and allowed to reach equilibrium.

The optical layout in Figure 17-20 shows that mirror *F* can take two positions. In position (1) it reflects the transmitted radiation onto the detector, in position (2) it reflects the radiation reflected from the sample onto the detector. When mirror *G* is raised into the converging radiation from mirror *P*, the radiation is reflected onto a horizontal mirror *H*, and then back onto *G* and *E* and hence to the detector. This allows the background transmission of the instrument to be measured under instrumental conditions as closely related to those for the reflection and transmission as are available without the use of a double-beam system.

Figure 17-21 illustrates the results obtained with a sample of Teflon, showing the reflection interferogram and the corresponding transmission and reflection transforms. It can be seen that the channel spectra from the transmitted to the reflected spectra are 180° out of phase, and that the interval from one maximum to the next changes with frequency. This is most noticeable in the reflection spectrum (curve B). With

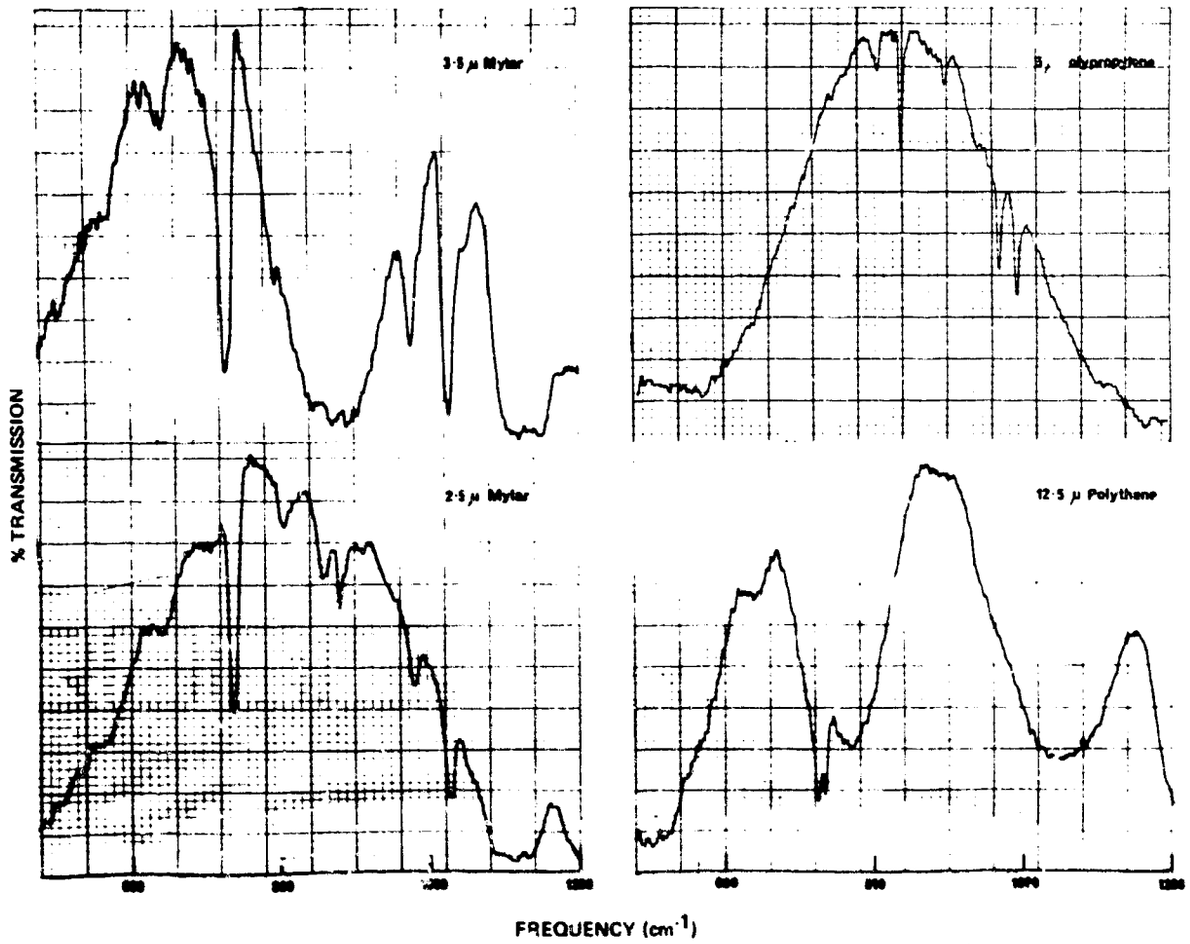


Figure 17-19. Comparison of Four Spectral Backgrounds Obtained With 3.5-μm and 2.5-μm Mylar, 5-μm Polypropylene, and 12.5-μm Polyethylene Using 8-μm Lowpass Filter (The transmission profile of the 12.5-μm polyethylene beam splitter is seen to be the most suitable for high frequency use.)

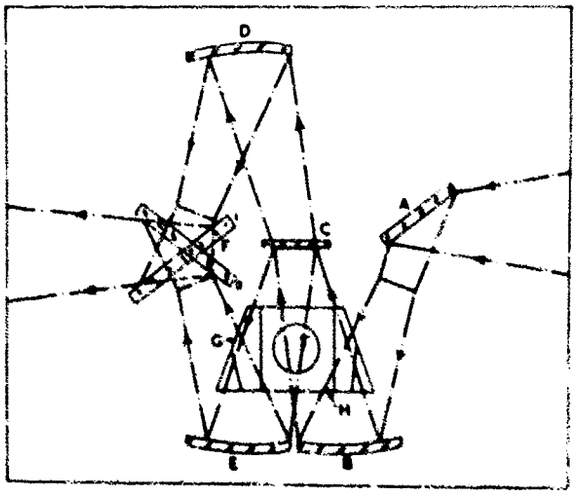


Figure 17-20. Optical Arrangement of Background, Reflection, and Transmission Unit

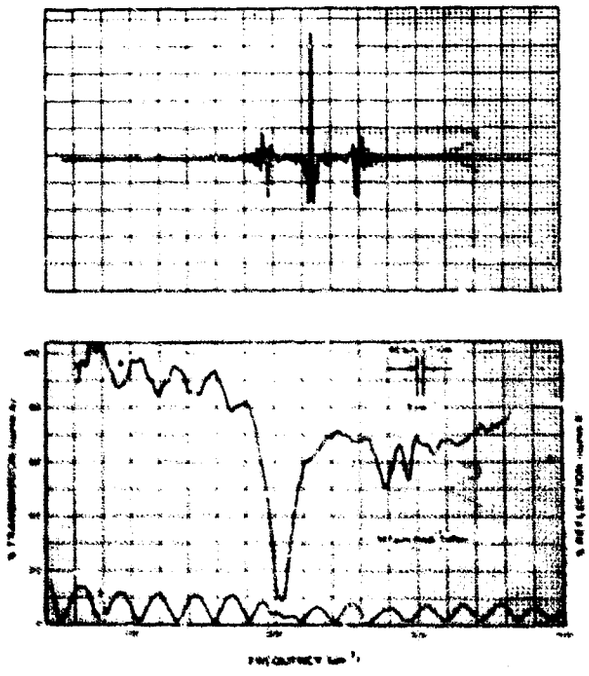


Figure 17-21. (*top*) Reflection Interferogram of Teflon (Maximum Optical Path Difference = 0.2 cm), and (*bottom*) Corresponding Ratioed Reflection and Transmission Spectra

the thickness of the material known, the refractive index can be estimated from:

$$\mu(\nu) = \frac{1}{2T\nu_{ch}} \qquad (17\text{-}7)$$

where ν_{ch} is the frequency interval from one maximum to the next of the channel spectra at the frequency ν. This method produces a refractive index function similar to the one in Figure 17-5 but the signal-to-noise ratio is considerably worse since accurate measurement of ν_{ch} from the output spectrum is quite difficult.

Figures 17-22 and 17-23, respectively, show the spectra and corresponding interferograms of a chromium complex and a cobalt complex.

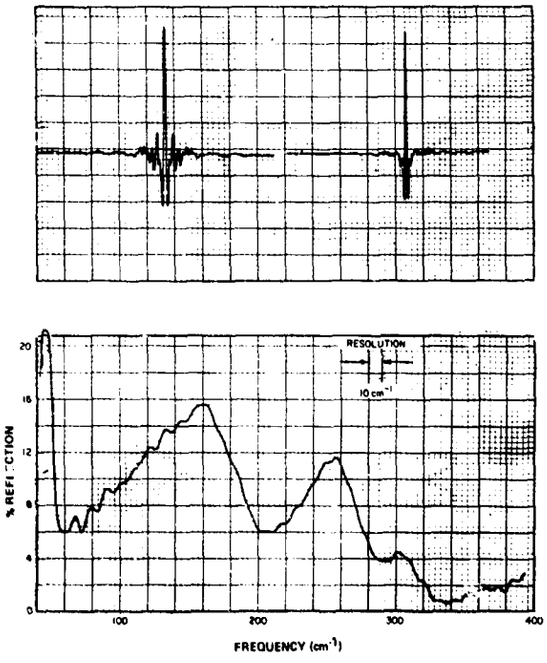


Figure 17-22. (*top*) Background and Reflection Interferograms of a Chromium Complex (Maximum Optical Path Difference = 0.1 cm), and (*bottom*) Corresponding Ratioed Reflection Spectrum

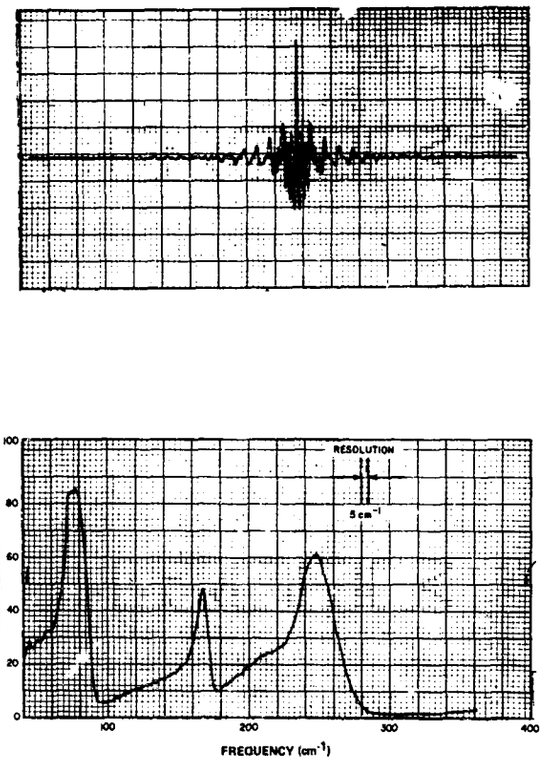


Figure 17-23. (*top*) Reflection Interferogram of a Cobalt Complex, E ⊥ C Axis (Maximum Optical Path Difference = 0.2 cm), and (*bottom*) Corresponding Ratioed Spectrum

17–6 CONCLUSIONS

The instrumentation and applications information contained in this paper are useful to chemical and physical spectroscopists using the Fourier method. Both the sample-handling and data-handling techniques are rapidly being improved, and the more these techniques become common practice in chemistry and physics laboratories of universities and industries, the more readily will fundamental problems be solved.

Acknowledgments

We are indebted to our colleagues at the R.I.I.C. Laboratories for their assistance during development of the instrumentation described here, and to all others involved in the production of this paper.

References

1. Thorpe, L. W., Milward, R. C., Hayward, G. C., and Yewin, J. D. (1969), presented at ICO Conference on Optical Instruments and Techniques, held at Reading, England, 14–19 July 1969, sponsored by the International Union on Pure and Applied Physics.

2. Martin, A. E. (1966) *Infrared Instrumentation and Techniques*, Elsevier, p. 74.

3. Chamberlain, J., Gibbs, J. E., and Gebbie, H. A. (1969) *Infrared Phys.* **9**: 185–209.

4. White, John U. (1942) *JOSA* **32**: 285.

5. Chantry, G. W., Evans, H. M., Chamberlain, J., and Gebbie, H. A. (1969) *Infrared Phys.* **9**: 85–93.

Contents

18. Technique for Determining Far Infrared Absorption Spectra of Small Samples of Liquids and Solids at Low Temperatures With a Fourier Spectrophotometer

James L. Lauer and Melvin E. Peterkin
Research and Development, Sun Oil Company
Marcus Hook, Pennsylvania

Abstract

A cone channel condenser is placed in the sample compartment of a RIIC FS-720 far infrared Michelson interferometer at the location of the source image. The concentrated radiation is then passed through the cooled sample and condensed onto the detector by a field lens and another cone channel condenser.

18–1 INTRODUCTION

When absorption spectra are to be obtained on small quantities of material, the use of radiation beam condensers has become standard practice among spectroscopists. Microscope optics are generally employed, usually consisting of complex mirror systems, to bring the radiation to an accessible position for sample alignment. The optical path through the sample is often also small, making it difficult to recognize weak absorption bands. Since weak absorption bands are the rule for most materials in the far infrared (the spectral region of primary interest to us) and the use of precise optical parts with samples at very low temperatures (the primary reason for undertaking this work) seemed unwarranted because of difficult alignment problems, the much simpler approach of conical radiation channel condensers was chosen.

Such cones have the advantages of easy introduction into and removal from the sample compartment of the interferometer available for this study (RIIC Model 720), and hence do not interfere with routine uses of the instrument. The angular spread of their exit beam makes it possible to obtain effectively longer optical paths with straight cells than collimated beams would allow. The circumstance that conical channel condensers are not imaging devices is no serious detriment, as the interferometer already includes such a cone channel ahead of the Golay detector. A simple plastic lens between the two channels, located at the exit from the sample compartment, ensures adequate transmission of the radiation to the detector.

18–2 GUIDELINES USED IN APPARATUS DESIGN

Figure 18-1 shows the optical layout of the RIIC Model 720 Far Infrared Interferometer.[1] Important features are the focal position within the sample compartment and the detector section consisting of a cone radiation channel condenser closed at its ends by a Golay cell and a black polyethylene lens, respectively.

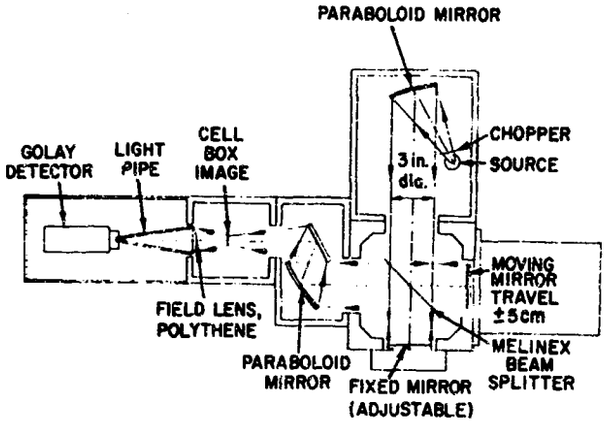


Figure 18-1. Ray Diagram for the RIIC Michelson Interferometer FS-720[1]

The lens serves several purposes. With its focal length properly chosen, it changes the divergence of all or most of the radiation impinging on it to fall within the range acceptable by the cone channel condenser. It is a limiting aperture, and it is a filter for visible and near infrared radiation. Also, the size of the source can be altered, providing for different image sizes and convergence of radiation within the sample compartment. Access to the compartment is provided by a removable lid.

Since our work required the determination of small changes in absorptivity with temperature, and the absorptivity was usually low, the largest source size was almost invariably required to get sufficient signal. Fortunately, the absorption bands of the liquids studied were generally broad and structureless so that the concomitant reduction of spectral resolution and wavenumber accuracy was not important. Hence, our designs and construction dealt exclusively with optimum signal conditions. The availability of large signals led us to choose the interferometer over a grating instrument in the first place. At maximum source size, the convergence or divergence of the extreme rays from the edge of the image in the sample compartment constituted a basic parameter for the design of our sample beam condensers.

For cooling samples to low temperatures and stabilizing them there, the sample cells had to be made attachable to the cold finger of a liquid nitrogen Dewar, with the connecting piece between cell and Dewar electrically heatable. Minimum size of sample cells was called for to achieve temperature uniformity. Preferably, there should be no direct physical contact but very close proximity between the cell and the cone channel condenser in order to transfer maximum amounts of radiation through the sample cell. Even with small cell volumes, temperature stability was difficult to maintain without elaborate temperature control devices so that relatively rapid scanning was important. Here, too, the interferometer proved invaluable because of its ability to survey a large wavelength region quickly.

The choice of a beam condenser system consisting of a cone channel condenser over one consisting of mirrors was made primarily on the basis of convenience, simplicity, and the absence of the need for an imaging device. A reduction of 4:1 in cross-sectional area of the radiation flux through the sample without loss of signal was entirely satisfactory.

18-3 OPTICAL CONSIDERATIONS

As was shown by Williamson[2] and Greenler[3], a cone channel condenser (CCC, in the form of a frustum of a cone), will only transmit radiation incident on its walls when the angle of divergence of the limiting ray with the cone axis is below a certain maximum value. This value depends essentially only on the ratio of the radii of the CCC and not on its length (assuming reflectivity is perfect, a condition generally true for metals in the far infrared). A converging lens at the entrance to the CCC can be helpful in bringing radiation within the required range. With a second CCC pointing in the same direction and placed in front of the lens, it is possible to pass just as much or even more radiation (if the first CCC constituted a limiting aperture) through the whole system, provided the location of the latter with respect to the focus of the lens and the general geometry are correct. The new CCC (CCC2) can be made to constitute an effective beam condenser for radiation detected at the small end of CCC1. Samples would be placed at or near the small end of CCC2. Another conical channel (CCC3) placed backwards immediately after CCC2 (assumed to collect the radiation first) can be helpful to reduce the divergence of the radiation impinging on the lens and on CCC1. However, it would have made reproducible sample location more difficult and was not needed.

Figure 18-2 shows how the trace of a limiting ray through CCC2, the beam condenser cone, can be used to determine its dimensions.

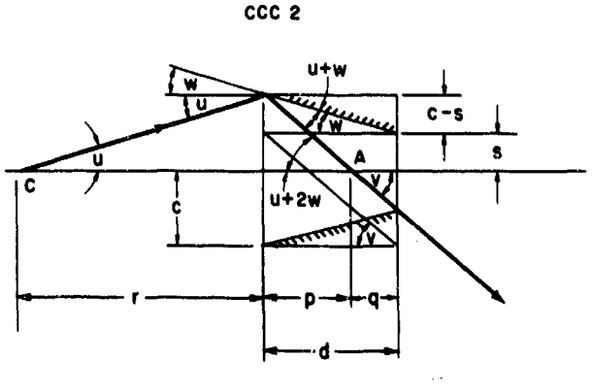


Figure 18-2. Ray Diagram for the Radiation Beam Condenser Used With Small Samples. Samples are placed adjacent to the small opening of CCC2

Let w be the half angle, c and s the large and small radii, and d the length of CCC2. Let u and v be the angles of the extreme rays with the optic axis. Then clearly

$$d=(c-s)\cot w \tag{18-1}$$

$$c+s=d\tan v. \tag{18-2}$$

Furthermore, from the law of optical reflection

$$v=u+2w. \tag{18-3}$$

From these relations the expression

$$s/c=(\cot w\tan(u+2w)-1)/(\cot w\tan(u+2w)+1) \tag{18-4}$$

can be derived for the effective convergence achieved by CCC2. It has been assumed, of course, that only one reflection occurs within CCC2. The analysis is easily extended to multiple reflections.

Figure 18-3 continues the analysis. The ray now

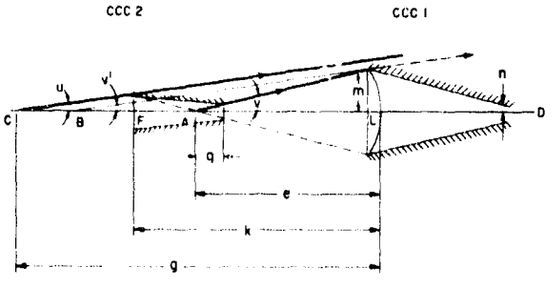


Figure 18-3. Ray Diagram for the Sample Chamber and Detector, Showing the Changes Produced by the Radiation Channel Condenser CCC2

shown as emerging from point A with angle v from the optic axis is the same critical ray that left CCC2 (for convenience, it has been reflected in a plane perpendicular to the paper and including the optic axis). It is assumed that A is to the right of F, the focus of lens L, so that the ray is still divergent from the optic axis (at angle v') after refraction by the lens L.

Using the nomenclature of Drude[4] for thin lenses, it follows that

$$(\tan v')/(\tan v)=e/e', \tag{18-5}$$

where e is the distance of an object A from the unit plane H (essentially the plane of the lens in our case), and e' is the distance of its image from the unit plane H′ (assumed here to coincide with H). Both e and e' are positive if A lies in front of (to the left of) H and B to the right of H′. Also,

$$(1/e)+(1/e')=(1/f) \tag{18-6}$$

(assuming equal focal lengths).

Combining Eqs. (18-5) and (18-6) results in

$$\tan v=(\tan v')f/(e-f). \tag{18-7}$$

Since e is positive and assumed smaller than f, $\tan v'$ is negative and so is e' $(=BL)$ and the ray is divergent even after passage through the lens L. As far as CCC1 is concerned, the ray might just as well have originated at B. Provided that

$$v'\leq \arcsin(n/m), \tag{18-8}$$

(the effective radius reduction possible[2,3] for CCC1), it will be transmitted to the detector D. Note that this ray might have originated at point C with a divergence angle u, and were it not for CCC2, it would have missed lens L and would not have been detected. Provided, therefore, that lens L reduces the effective aperture of the instrument below that of the optics ahead, it is possible to gain signal strength by the insertion of CCC2. This indeed was our situation. There is no observable deterioration of resolution and wavenumber accuracy under our operating conditions.

The location of CCC2 with respect to L can be defined in the following way. Refer to Figures 18-2 and 18-3. The angle u of the limiting ray with the

optic axis is assumed known. Angle w of CCC2 and all the parameters pertaining to CCC1 are also assumed known. Then, by trigonometry

$$q = s \cot v = s \cot (u+2w) \tag{18-9}$$

$$e = m \cot (u+2w), \tag{18-10}$$

and

$$k = e - q + d = (m-s) \cot (u+2w) + (c-s) \cot w. \tag{18-11}$$

In practice, the CCC2 angle w is the smallest angle conveniently cut on the lathe and, knowing u, d, c, and s, the other parameters pertaining to the cone channel condenser CCC2 are then determined. However, g, the distance from the point (C) where the peripheral ray intersects the optic axis at angle u to the lens, must also be known if the actual values of the radii c and s of CCC2 are to be calculated and not just their ratio. This distance, and the effective angular opening, change as the interferometer mirror is moved from the position of zero path difference, as is well known[5], but a compromise is easily reached. This variation is, in turn, reflected as an increase of relative signal strength with increasing wavenumber when the Fourier transformation is carried out, but since absorption spectra are invariably ratioed to a reference, this effect is not detrimental. The associated wavenumber shift can be corrected.

By combining Eqs. (18-1) and (18-2), the following useful relations are obtained, namely,

$$c = d (\tan v + \tan w)/2, \tag{18-12}$$

and

$$s = d (\tan v - \tan w)/2, \tag{18-13}$$

which can be combined with the relation for g,

$$g - e + q - d = c \cot u, \tag{18-14}$$

to give

$$d = 2(g-e)/(1+\tan v \cot u + \tan w \cot u + \cot v \tan w). \tag{18-15}$$

Now all the required parameters are determined. As an illustration, $u = 14$ deg, $w = 8$ deg, $m = 2.1$ cm, $n = 0.15$ cm and $g = 5.8$ cm. From Eq. (18-8), v' (max) $\simeq 4$ deg or $\tan v' = -0.071$. Since

$$v = u + 2w = 30 \text{ deg} \quad \text{or} \quad \tan v = 0.58,$$

Eq. (18-7) gave $(e/f) = 0.88$. Equation (18-10) then yielded the values of $e = 3.64$ cm and $f = 4.14$ cm. Equations (18-12) and (18-13) resulted in $c = 0.359\,d$ and $s = 0.218\,d$. The value of d was found from Eq. (18-15) as $d = 1.06$ cm, thus fixing c and s as $c = 0.39$ and $s = 0.23$ cm. Since the radius of the image of the source in the sample chamber as provided by the instrument maker was 0.6 cm for the source diameter used, a reduction of cross-sectional area required for the sample of almost 7:1 was achievable. From Eq. (18-11) the position of CCC2 is given by $k = 4.37$ cm.

18-4 APPARATUS

Figure 18-4 shows the two-cone channel condensers used most frequently in this work. They are supported on a pedestal whose height can be accurately adjusted and are located in the sample compartment by means of reference marks. They were made out of brass and polished to the finest finish possible along their interior surfaces. The entrance and exit diameters are 2.5 cm and 0.5 cm, respectively. The cone angles are 16 deg (half angle 8 deg), and the overall length of the two-cone channel condensers is about 8 cm. Before use, the brass was gold plated to a mirror finish to avoid tarnish. One of these condensers was used primarily with the cylindrical cell directly attached to it while the other one was used in low-temperature work. For the latter purpose, small cylindrical cells were attached to the copper strap suspended from the cold finger of the liquid nitrogen Dewar* shown in Figure 18-5. The same copper strap also holds two electric heaters which allow for temperature stabilization by reducing the heat flow from the cell. Cells of different lengths are used depending on the characteristics of the sample. The temperatures of the sample cells are determined by copper-constantan thermocouples directly attached to the

* Manufactured by Huffman Division of Minnesota Valley Engineering Company.

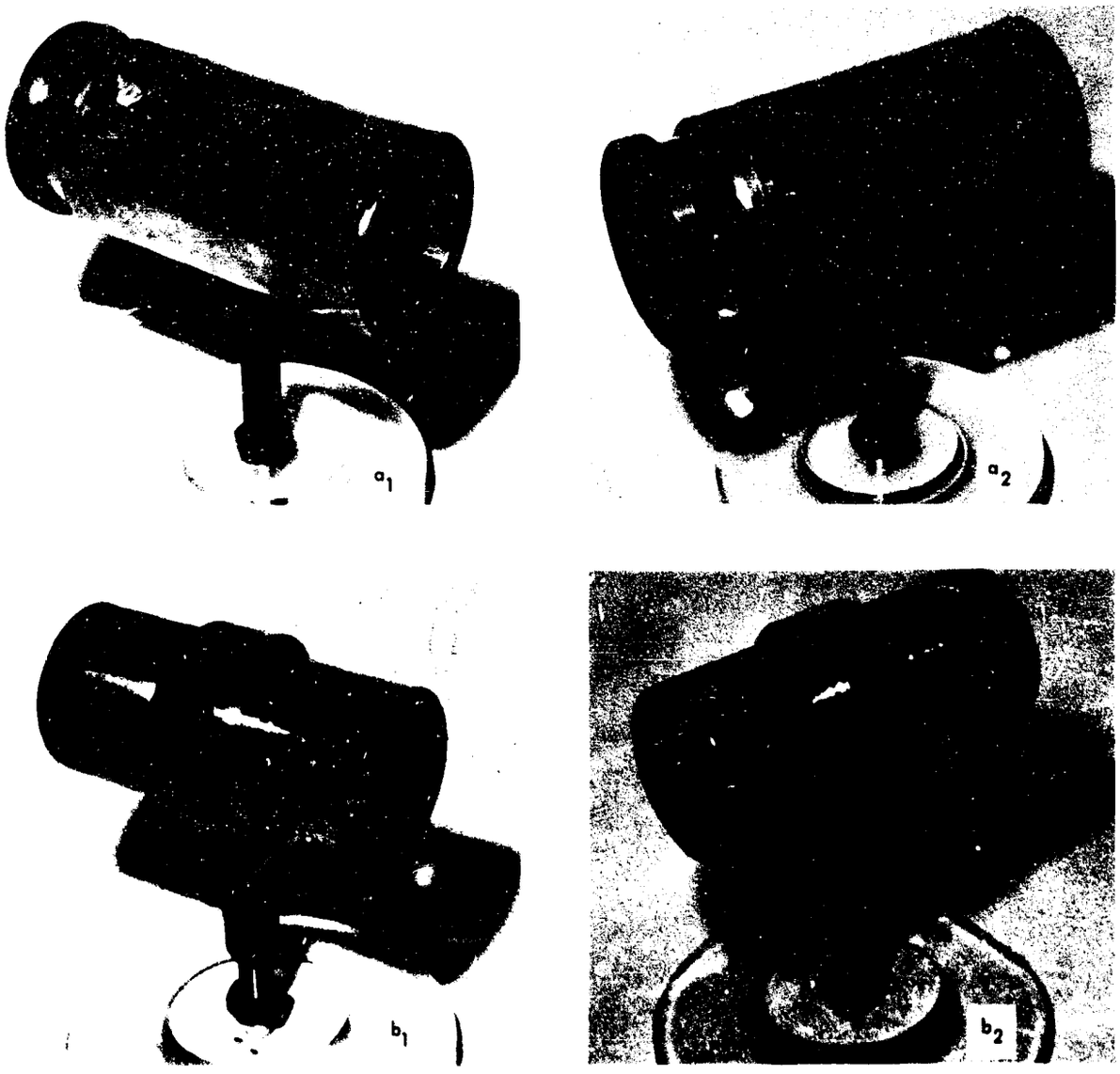


Figure 18-4. Radiation Beam Condensers. a) With cylindrical liquid cell attached. b) With attachment, for use with cooled sample cells

cell bodies with a special indium solder. The cell windows are made of high density polyethylene† selected for high transmission in the far infrared. They are attached to the cell body by caps screwed over the ends. At the ends of the cell body, O-rings seal the cell contents against the cell windows. The Dewar is a specially designed liquid nitrogen container with double walls made out of stainless steel, and the vacuum around the inner part holding the cold finger is provided by the sample compartment of the instrument itself. The Dewar assembly includes a duplicate lid which takes the place of the usual sample compartment lid. With Viton O-rings it was possible to maintain sealing of liquids down to temperatures of minus 150 deg C. It was preferable, however, to fill the sample compartment with dry nitrogen when allowing the Dewar flask and its attachments to return to room temperature to prevent leakage past the O-rings. Temperatures intermediate between room temperature and that of liquid nitrogen could be reached and maintained for reasonable lengths of time by adjusting the current through the heaters with a variable voltage transformer. Because of this rather crude, though effective, method of temperature control, it was necessary to constantly check that the liquid nitrogen supply in the Dewar flask was maintained.

† Made by Chemische Werke Hüls, "Vestolen A."

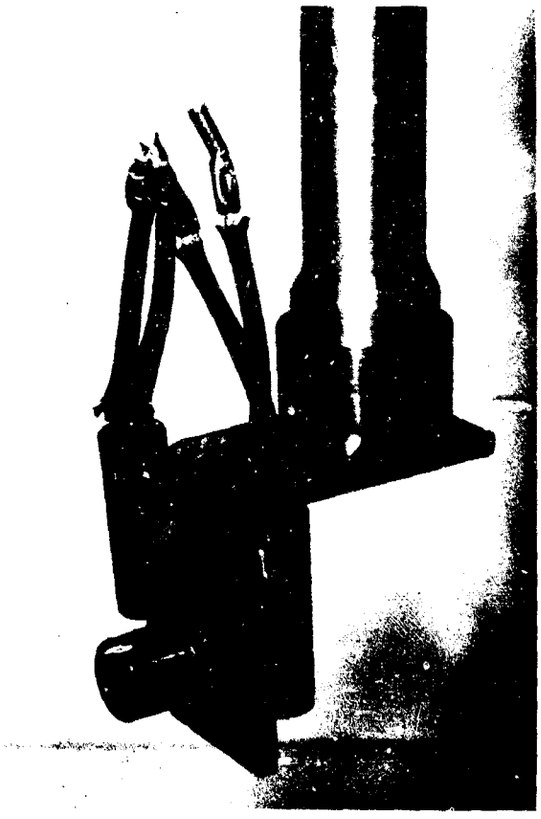


Figure 18-5. Dewar Attachment With Liquid Cell for Low Temperature Spectroscopy in the Far Infrared

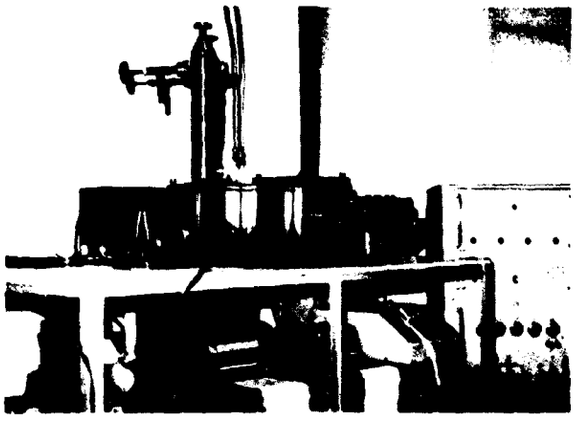


Figure 18-6. Far Infrared Fourier Interferometer Installation for Low Temperature Work

The entire far infrared installation with the Dewar flask in place is shown in Figure 18-6. All the output data were produced on punched paper tape by the punch located below the optical parts, and the Fourier transformations to spectra were carried out with the help of a teletypewriter communicating over telephone wires with a G.E. 635 Computer at company headquarters. The interferometer shown here was the first of its kind providing BCD output produced by the Research and Industrial Instruments Company (RIIC) of London. Experiences with it stimulated many developments included in later models by RIIC.

It was found possible to maintain temperatures for a sufficiently long time to obtain spectra with a theoretical resolution of about 2.5 cm^{-1} (about 30 min). The apparatus is used not only for cooled liquids but for solids as well. Solids, dispersed in high density polyethylene and molded in the form of thin discs, are supported very conveniently at the exit end of the cone channel condenser (CCC2).

18–5 PERFORMANCE

Figure 18-7 shows the ratio of signal with cone

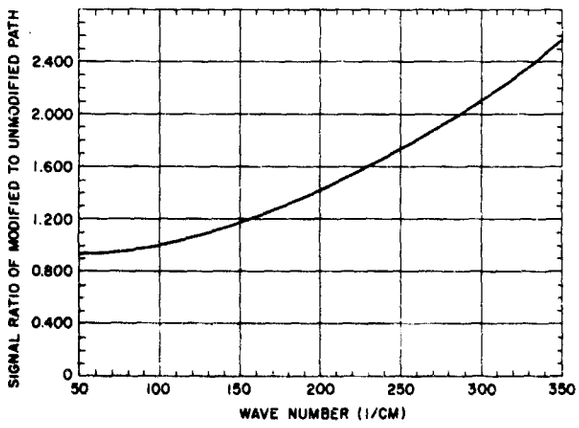


Figure 18-7. Ratio of Detector Signal Obtained With and Without Cone Channel Condenser. The entire instrument is evacuated

channel condenser (CCC2) to that without over a wide wavenumber region. The gain attributable to CCC2 is more pronounced with the higher wavenumber. The relationship is nearly linear. As already mentioned, the results in Figure 18-7 are caused by the increase of solid angle obtained at the entrance to the detector compartment. This increase in signal strength is especially valuable because of the reduction of flux diameter through the sample, which otherwise would be accompanied by a signal loss. A deterioration of resolution and wavenumber accuracy was anticipated. However, since normal runs used only about 1000 mirror positions corresponding to a theoretical resolution of 2.5 cm^{-1}, no significant differences were observed, as can be seen by a comparison of the water spectra obtained both with and without CCC in the sample compartment (Figures 18-8 and 18-9). The radiation path through the sample com-

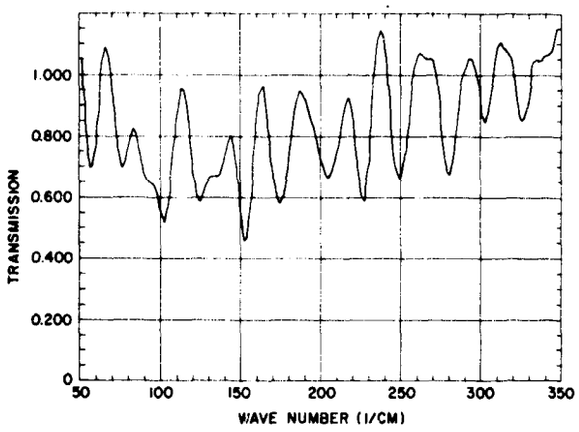


Figure 18-8. Water Vapor Spectrum Obtained With the Cone Channel Condenser in the Sample Compartment Exposed to the Atmosphere. The interferometer was evacuated. This spectrum was smoothed (for comparison with Figure 18-9) to show the main bands only

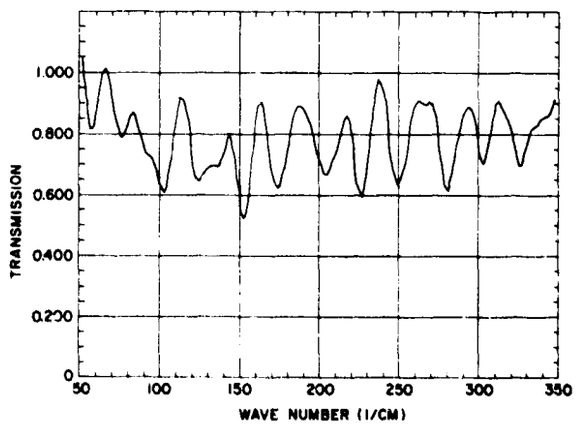


Figure 18-9. Water Vapor Spectrum Obtained in the Same Way as That of Figure 18-8, except for the Absence of the Cone Channel Condenser

partment is somewhat greater in the former case. If necessary, the digitized output and the characteristics of the interferometer can be used to a large extent to correct any errors thus introduced, for it is possible—most simply by empirical calibration based on the water spectra—to bring a spectrum back to standard conditions by computational manipulations. Since this work did not require great accuracy, these corrections were unnecessary.

Because of the increased angle of the radiation through the sample cells over the angles for which the instrument had originally been designed, there is some increase in effective radiation path lengths through the samples when the cone channel condenser is used. This increase makes it possible to obtain sufficient absorption of far-infrared radiation with smaller amounts of fluid. However, absorbance and cell length are still proportional.

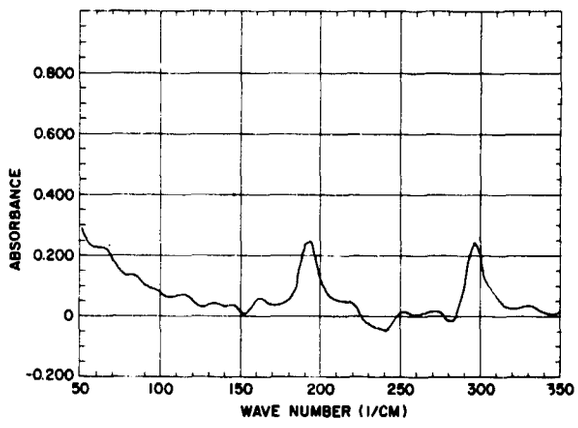


Figure 18-10. Spectrum of a 4 Percent Chlorobenzene in Cyclohexane Solution Obtained in the 1.2 Cm Cold Cell Placed Behind CCC2. The reference is the solvent

Figure 18-10 shows a spectrum obtained with a 4 percent solution of chlorobenzene in cyclohexane (referenced to cyclohexane) when the cell attached to the Dewar flask is used in conjunction with one of the cone channel condensers. In this case, the radiation emerging from the cone channel condenser has to jump a small distance before it can enter the sample cell. Nevertheless, good signal strength is still obtained and the spectrum is still of good quality even though there is some bouncing of radiation at the sample cell walls. It should be realized that a good part of the radiation transmitted by the sample cell has not been affected by the cone channel condenser at all; this central part of the radiation has traversed the sample directly. There is not much difference between the spectrum shown in Figure 18-10 and the one obtained in the manner specified by the instrument maker, that is, without cone channel condenser (Figure 18-12). Figure 18-11 shows essentially the

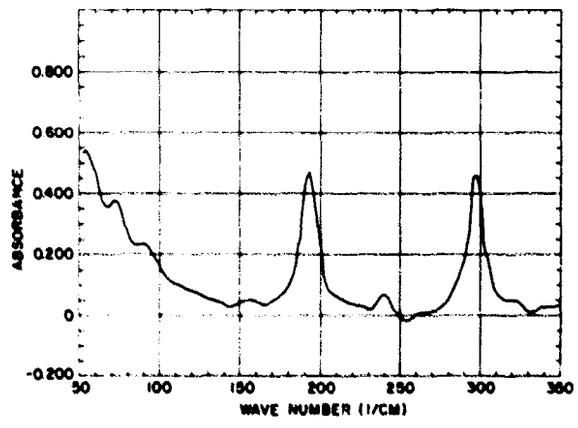


Figure 18-11. Spectrum of a 5 Percent Chlorobenzene in Cyclohexane Solution Obtained in the 1.2 Cm Cylindrical Cell Directly Attached to a Cone Channel Condenser. The reference is the pure solvent

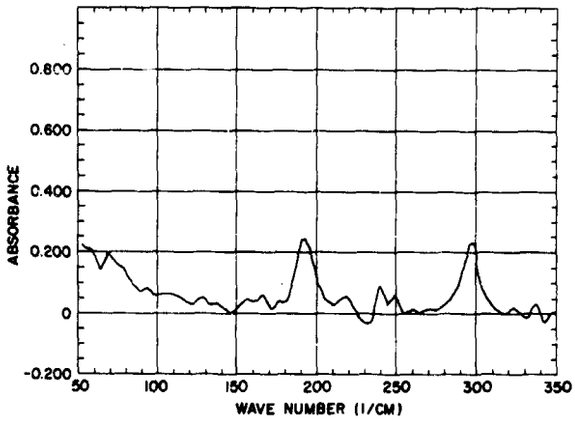


Figure 18-12. Spectrum of a 4 Percent Chlorobenzene in Cyclohexane Solution Obtained in a Standard 1 Cm Long Liquid Cell. No cone channel condenser was used. The reference is the pure solvent

same solution spectrum obtained with the liquid cell directly attached to the cone channel condenser. Clearly, the effective absorption path has almost been doubled because of bouncing of radiation at the cell walls. If necessary, the small "shading toward the visible" produced by the CCC2 can be corrected by applying the necessary wavenumber shift.

The very significant bending of the radiation that occurs in the sample cells when the CCC2 has been used is shown clearly by the spectrum of Figure 18-13. This figure contains absorbance data of the same chlorobenzene solution, but this time referred to a vacuum background (with the CCC2 and all other conditions being equal). Hence, there is a peak at 240 cm^{-1}, which is caused by the cyclohexane solvent. Note that the background continuously rises with

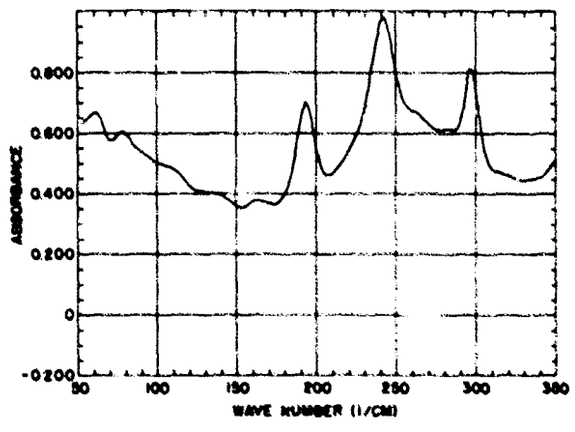


Figure 18-13. Spectrum of a 4 Percent Chlorobenzene in Cyclohexane Solution Obtained in the 1.2 Cm Cylindrical Cell Directly Attached to a Cone Channel Condenser. The reference is the evacuated chamber with the cone channel condenser

wavenumber, as would be expected from the increase of refractive index with increased wavenumber. Such data can be helpful for crude estimates of the index of refraction of liquids. In general, spectra of solutions should always be ratioed to those of the pure solvent.

An illustration of the changes observed in the far-infrared spectra of some liquids with a decrease in temperature is shown in Figure 18-14. The reference for these absorbances is the room temperature spectrum, so as to emphasize the differences. Notice that the spectrum of methylcyclohexane changes considerably as the temperature is reduced, especially after solidification (at the lowest temperature shown), but already before. Changes of far-infrared absorption in the liquid phase have been ascribed by various

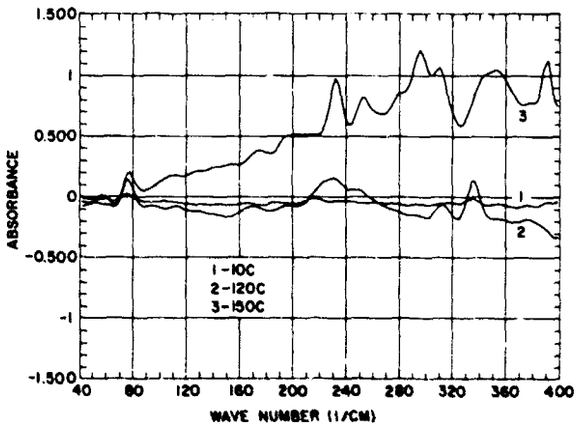


Figure 18-14. Changes in the Far Infrared Spectrum of Methylcyclohexane with Temperature. The reference is methylcyclohexane at room temperature. Spectrum 3 refers to the solid

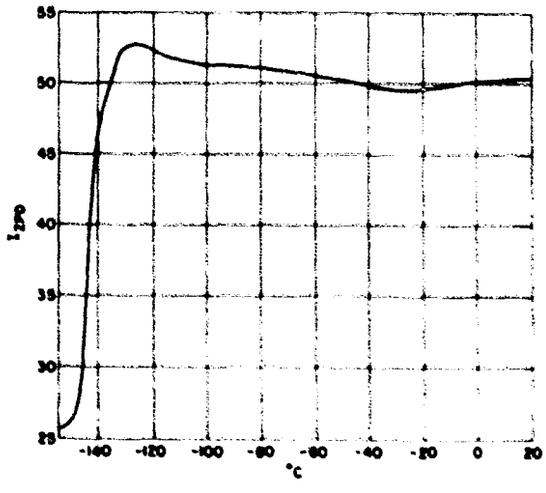


Figure 18-15. Variation of Detector Signal at Zero Path Difference With the Temperature of Methylcyclohexane in the Optical Path. A weighted transmission average over the 40-400 cm^{-1} range is represented

investigators to changes of rotational freedom, pseudo-lattice transitions, and changes in the Debye absorption.[6,7,8] It is not the purpose of this paper to discuss these changes in detail, but it is of interest to show how interferometry in the far infrared can be used to emphasize them. Figure 18-15 contains a plot of the detector signal against temperature, when the movable interferometer mirror is at zero path difference, again for methylcyclohexane. The wavenumber range was limited to 40 to 400 cm^{-1}. Accordingly, an average over the entire range (though weighted by the instrumental parameters) is reflected by this signal. As the temperature is lowered, the signal first increases and then abruptly decreases as the sample is solidified. In some materials more complex changes have been noticed. A discussion of these effects and the individual spectra they reflect will be presented in a different publication.

Acknowledgment

The authors would like to thank Mr. Thomas A. Smith of the Sun Oil Research and Development Machine Shop for his excellent work in the construction of the cone channel condensers. Without his diligence and meticulous attention to details, this work would not have been possible. The authors are grateful to the Sun Oil Company for their support of this work and permission to publish it.

References

1. Milward, R. C. (1969) Molecular spectroscopy, *Proc. Conf. Inst. of Petroleum* (*London* 17–19 April 1968), P. Hepple, Ed., held at Brighton, England. Published (1969) by Elsevier, Amsterdam. pp. 81–96.

2. Williamson, D. E. (1952) *J. Opt. Soc. Am.* **42**: 712.

3. Greenler, R. G. (1967) *J. Opt. Soc. Am.* **57**: 1062.

4. Drude, P. (1939) *Theory of Optics*, Longmans, Green and Co., New York. pp. 14ff.

5. Vanasse, G. A., and Sakai, H. (1967) *Progress in Optics, VI*, E. Wolf, Ed., North-Holland, Amsterdam. pp. 261ff.

6. Chantry, G. W., Gebbie, H. A., Lassier, B. and Wyllie, G. (1967) *Nature* **214**: 165.

7. Ewing, G. E. (1969) *Accounts Chem. Res.* **2**: 168.

8. Hill, N. E. (1968) *Chem. Phys. Letters* **2**: 7.

Contents

19. An Interferometer with Spherical Mirrors for Fourier Spectroscopy

M. Botten a
Ball Brothers Research Corporation
Boulder, Colorado
and
H. J. Bolle
Meteorological Institute of the University of Munich
Munich, Germany

Abstract

The interferometer consists of a beamsplitter and two spherical mirrors and is intended for application in planetary spectroscopy from sounding rockets. Each mirror images the source on the detector. The usable path difference is twice the focal tolerance. The transfer function is derived, and results obtained with a laboratory model are presented.

19–1 INTRODUCTION

For the study of planetary atmospheres in the far infrared, the use of a rocket-borne telescope combined with a Fourier spectrometer and a liquid-helium cooled detector was considered. The principal subject of this study is the emission spectrum of Jupiter. The peak is anticipated to lie in a region of strong NH_3 absorption between $100\,cm^{-1}$ and $300\,cm^{-1}$. A calculated spectrum of a model atmosphere containing only ammonia as a radiating gas above the cloud layer is shown in Figure 19-1. A cloud-top temperature of 168° K is assumed. The NH_3 spectrum consists of groups of unresolved lines, which appear as bands with a half-intensity width of about $5\,cm^{-1}$, spaced at intervals of about $20\,cm^{-1}$. The anticipated average spectral flux is of the order of $10^{-6}\,W\,cm^{-2}\,sr^{-1}\,cm^{+1}$.

To detect the ammonia bands and to make maximum use of the available flux in the limited time of the

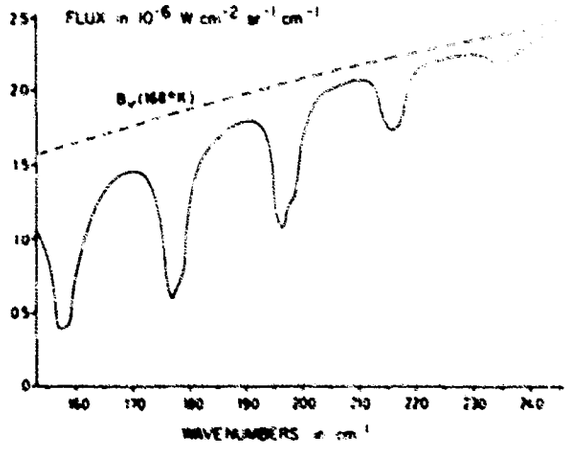


Figure 19-1. Calculated Spectrum of the Jupiter Atmosphere Between 150 cm⁻¹ and 250 cm⁻¹

rocket flight, it was decided to use a simple interferometer as shown in Figure 19-2. It consists of a

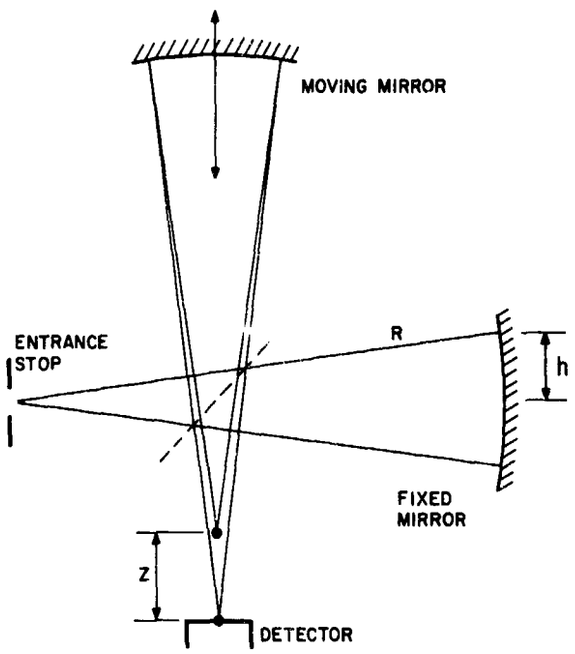


Figure 19-2. Spherical Mirror Interferometer

mylar beamsplitter and two spherical mirrors, one of which is movable. The planetary image formed by the telescope is placed at the center of curvature of the fixed mirror, and the detector is placed at its geometrical image. The two images formed by the fixed and the movable mirror are separated axially by a distance equal to the path-difference z in the interferometer.

The image separation evidently places an upper limit on the usable path-difference, but, as will be shown in the following, the spectral resolution necessary for detection of the NH_3 bands can easily be achieved. The proposed interferometer is easy to match to the telescope and has the advantage of greater simplicity, compared to the conventional arrangement. This is of special interest in rocket applications, where weight and size are important design factors.

19-2 PROPERTIES OF THE INTERFEROMETER

The properties of the interferometer can best be assessed by evaluation of the transfer function S. Consider the planetary image as a circular source of uniform brightness and radius r, and let the interferometer beam be limited by a circular aperture stop of radius h at the fixed mirror. According to Steel,[1] the transfer function for monochromatic radiation of wavelength λ is then given by

$$S=(2/r^2)(2/h^2)\int_0^r x\,dx\int_0^h y\,dy$$

$$\exp\{-\tfrac{1}{2}ikz(x^2+y^2)/R^2\}J_0(zkxy/R^2), \quad (19\text{-}1)$$

where x and y are radial variables in the source and the interferometer aperture stop, respectively; P is the radius of curvature of the mirrors, z is the path difference and $k=2\pi/\lambda$.

Consider first a point source. Equation (19-1) then reduces to

$$S=(2/h^2)\int_0^h y\,dy\,\exp\{-\tfrac{1}{2}ikz(y/R)^2\}$$

$$=\exp(-\tfrac{1}{2}iu)\sin\tfrac{1}{2}u/(\tfrac{1}{2}u) \quad (19\text{-}2)$$

with

$$u=\tfrac{1}{2}kz(h/R)^2. \quad (19\text{-}3)$$

This function is of the same form as the transfer function for the Twyman-Green interferometer if the latter were used with a circular source of radius h and a collimator with a focal length R. Analagous to the Twyman-Green interferometer, the maximum permitted path-difference z_{max} is defined as the one where the modulation in the interferogram has decreased to 64 per cent.[2] This occurs when $u=\pi$ and gives

$$z_{max}=\pm\lambda(R/h)^2. \quad (19\text{-}4)$$

It is of interest to note that the function in Eq. (19-2) is identical to the one describing the on-axis diffracted light amplitude in the diffraction image formed by the movable mirror at the detector. The conventional focal tolerance is defined as the distance where the on-axis irradiance has decreased to 80 percent of the value at the geometrical image.[3] In Eq. (19-2) this corresponds to $u=\pi/2$. Hence, the simple conclusion is that the permitted maximum path-difference is twice the conventional focal tolerance in the image, formed by either mirror, of its center of curvature.

An alternative interpretation of this result can be given by considering the superposition of the diffraction images of the point source formed by the two mirrors. As is well known, within the range defined by Eq. (19-4), the Fresnel diffraction pattern

formed by the movable mirror is very similar in structure to the Airy pattern formed by the fixed mirror. Hence, high modulation upon variation of z can be expected. In fact, this representation can be used for an alternative derivation of the transfer function. Since this requires a rather lengthy mathematical derivation, without presenting essential new views, no proof will be given here.

The spectral resolution in wavenumbers $\Delta\sigma$ is half the reciprocal of the maximum path-difference.[2] Hence, the resolving power becomes

$$\sigma/\Delta\sigma = 2(R/h)^2. \tag{19-5}$$

If the interferometer is properly matched to the telescope, this is a function only of the f/number of the latter. With an $f/6$ telescope, for instance, a resolving power of 288 can be achieved. In the wavelength region of interest, this is amply sufficient to identify the NH_3 bands with certainty. The maximum permitted path-difference at a wavelength of 30 μm is ± 4.3 mm. In either branch of the interferogram up to nine signatures corresponding to the band spacing should be detectable.

So far, only a point source has been considered. For practical applications it is necessary to take the finite size of the source into account. For the transfer function, return to Eq. (19-1). A closed form solution has been given by Steel[1] for $r = h$, but is not available for the general case. For the present purpose an adequate approximation can be found by expansion of the integrand in a power series of x. For small values of r,

$$S = \exp(-\tfrac{1}{2} iu)\{(\sin \tfrac{1}{2} u)/(\tfrac{1}{2} u) - \tfrac{1}{2} iw \exp(-\tfrac{1}{2} iw)\} \tag{19-6}$$

where $w = \frac{1}{2} kz(r/R)^2$.

Using a telescope with a focal length of 210 cm, for instance, the radius of the Jupiter image is about 0.2 mm. Hence, at maximum path-difference $w = \pi(r/h)^2$, which can be made negligibly small by choosing h sufficiently large, for instance $h = 4$ mm. For the reduction of the interferogram the conventional transfer function can then be applied.

19-3 SIMULATION OF THE INTERFEROMETER PERFORMANCE

A preliminary experiment was made to demonstrate the performance of the proposed instrument. For this purpose, an interferometer was set up which could accept an $f/6$ beam as planned for the rocket telescope. The mirror radius is 300 mm, the diameter 50 mm. A mercury lamp is used as a source, modulated at 15 Hz. The detector is a Golay cell with a quartz window. Additional filtering was accomplished by a 0.1-mm-thick sheet of black polyethylene. The shortest wavelength which then can be measured is about 45 μm. The field stop is defined by the size of the detector and has a radius of 1.5 mm. Hence, the maximum value of w is 0.011, which is negligibly small. According to Eq. (19-4), the maximum permitted path-difference is ± 6.4 mm. However, the interferometer was operated with a maximum path-difference of only ± 3 mm, this being the limit imposed by the drive system.

An interferogram obtained by running the interferometer in normal laboratory air and the computed spectrum are shown in Figure 19-3. The upper part

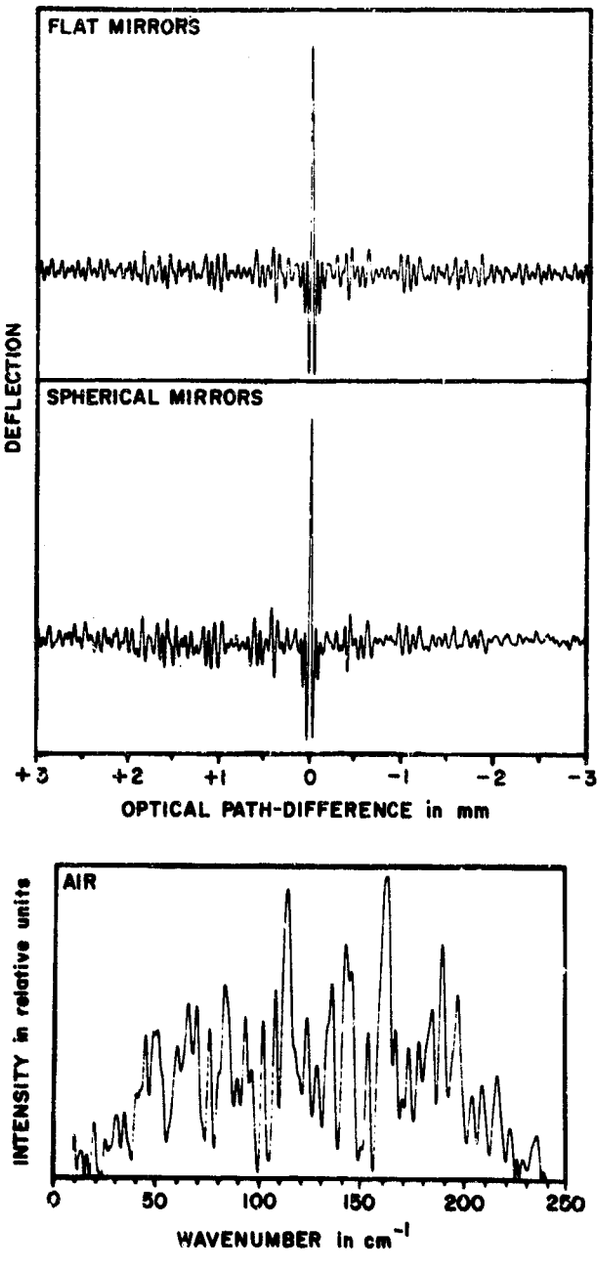


Figure 19-3. Interferogram and Spectrum Obtained With a Spherical Mirror Interferometer in the Laboratory

of the figure shows a comparison between interferograms obtained with flat and with spherical mirrors. The lower part of the figure gives a spectrum of 170 cm laboratory air derived from an interferogram obtained with spherical mirrors. There is a noticeable asymmetry, the cause of which has not yet been fully identified, but is probably a minor technical imperfection of the mirror drive. A triangular apodization function was used in the complex transform of 1173 sampling points at 2.5 μ distance. The absorption features are all due to water vapor in the 1.5-m air path. The details in the spectrum indicate a spectral resolution close to the theoretically achievable value.

References

1. Steel, W. H. (1964) *J. Opt. Soc. Am* **54**: 151.

2. Vanasse, G. A., and Sakai, H. (1967) *Progress in Optics*, *VI*. E. Wolf, Ed., North-Holland, Amsterdam, p. 282.

3. Born, M., and Wolf, E. (1959) *Principles of Optics*, Pergamon Press, New York, p. 440.

Contents

20. Performance of a Simple Spherical Lamellar Grating Beamsplitter by Wavefront Division

J. P. Hansen and John Strong
Astronomy Research Facility
University of Massachusetts
Amherst, Massachusetts

Abstract

A simple lamellar grating Fourier spectrometer for the 3 to 100 cm^{-1} region with resolution between 0.4 and 0.5 cm^{-1} and its performance are described. Spectra are given of atmospheric water vapor absorption and HA3 mercury lamp emission in both a vacuum and a helium atmosphere, as well as channel spectra of quartz plates.

20–1 THE INSTRUMENT

The Fourier spectrometer described involves a minimum of components owing to the fact that the lamellar grating provides three necessary functions: path difference, wavefront division, and focusing, which avoid the need for collimators and beamsplitters, such as are employed in other interferometers. Figure 20-1 illustrates the main components of the interferometer. These components are: a source; a chopper; the grating, and a detector. With the exception of a black paper filter placed at the entrance aperture, and a cold quartz focusing lens immediately in front of the germanium chip (detector), there are no dispersive elements in the path of the radiation. The filter essentially eliminates high-frequency radiation; it can be considered a linear element in the frequency range of interest, 3 to 100 cm^{-1}. The complete interferometer is contained in an evacuable chamber. The interferometer chamber can be evacuated to pressures below 10 microns, and various gases can be admitted for absorption studies. The grating and detector end of the chamber has been designed so that it can be mounted on a 300-ft absorption cell that is available to us, and with which we plan to conduct further investigations. In addition, there is the necessary electronic system for the processing of response and for display.

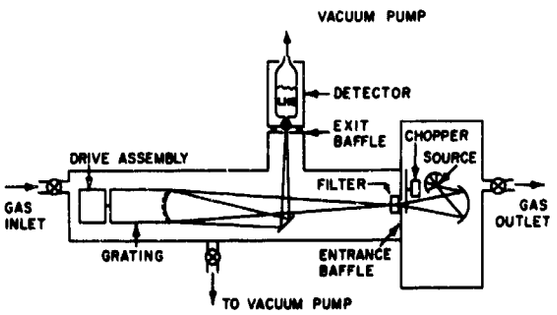


Figure 20-1. Schematic of Basic Interferometer

The grating facets are made of steel. At first, the reflecting surfaces on them were flats. The

present spherical surfaces were generated with the moving facets held rigidly in place (with zero path difference). Then a spherical surface ($R = 152$ cm) was ground and polished on the grating, to replace the flat surface. The finished and mounted grating is shown diagrammatically in Figure 20-2. In order to

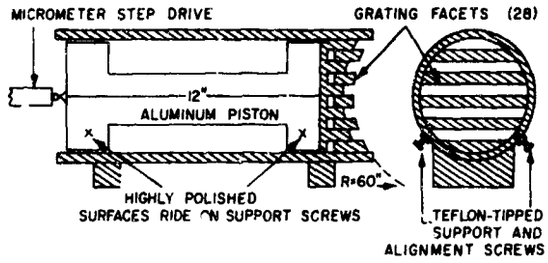


Figure 20-2. Assembled Grating: Front and Side View

assure parallel motion of the moving facets with respect to the stationary ones, the moving facets, attached to a 30-cm aluminum piston, are defined by and slide on four Teflon-tipped alignment and support screws, adjusted to make the foci of the moving and the stationary facets coincide for zero-path difference; displacing them along the center line of the grating as path difference increases. Grating advancement takes place through a micrometer screw that pushes on a bearing ball mounted at the center of the piston. This particular arrangement assures parallel motion of the grating, as the direction of the moving piston (not rigidly fastened to the drive mechanism) is constrained by the alignment screws. When data are being recorded, the grating path difference is stepped back from its maximum through zero, then the movable part of the grating is returned to maximum path difference at a fast rate by means of a return mechanism. Both synchronous and variable speed motors are used in conjunction with a Geneva movement to advance the grating. Step size, grating speed, and maximum path difference are adjustable to allow for the minimum recording time and specified resolution that is appropriate for various spectra. A slide wire potentiometer monitors grating position at all times.

A helium-cooled germanium bolometer (provided by Infrared Laboratories, Tucson, Arizona) is used as the detector. The operating temperature of 1.9° K is achieved by pumping on the helium with a high capacity vacuum pump. The noise equivalent power has been measured and calculated to be 2×10^{-13} watts. This value is optimized by proper selection of a bias battery.

The detector output is directly applied to a low noise preamplifier and then to a lock-in amplifier which is referenced to the 100-Hz synchronous chopper placed in front of the input aperture (Figure 20-3). The amplified detector output as a function of

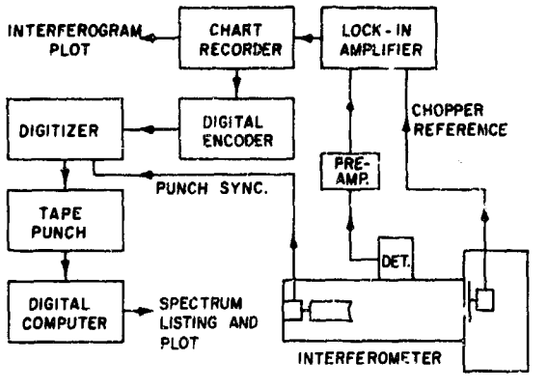


Figure 20-3. Block Diagram of Fourier Spectrometer

path difference—the interferogram—is recorded as a plot on a strip chart recorder. The output is simultaneously digitized using a linear encoder whose output is directly proportional to the recorder pen position. The data are coded and punched on paper tape for feeding to the computer.

20-2 DESIGN CONSIDERATIONS

The multiplex and throughput advantages of Fourier spectroscopy are well described in the literature.[1-3] Here we discuss their realization by a particular instrument. Our spherical lamellar[1] grating was designed for the 3 to 100 cm^{-1} region. The grating diameter is 9 cm and the radius of curvature 152 cm. The grating constant is 0.635 cm. The reflected energy oscillates between the zeroth and higher orders as the groove depth is varied.[1] The entrance and exit apertures are identical; they were chosen in such a way that first- and higher-order radiation of the highest frequency present (100 cm^{-1}) does not reach the detector, in order to yield optimum modulation. The first-order grating equation is

$$\lambda_c = a \sin \alpha, \tag{20-1}$$

where λ_c is the shortest wavelength present in the radiation to be analyzed, a is the grating constant, and α the angle of incidence of the radiation on the grating. Since the grating is operated at near normal incidence, $\sin \alpha \approx \alpha \approx \frac{W}{R}$ where W is the slit width and R the radius of curvature of the grating. Thus for $\sigma_c = 100$ cm^{-1}, the maximum slit width is 2.4 cm. Actually, our entrance aperture is a round hole of 2.2 cm diameter; the active aperture of our detector at the focus of the sphere is the same.

The low frequency limit of the interferometer is mainly limited by the grating constant. When the wavelength of the radiation incident on the grating approaches the size of the grating constant, waveguide effects occur, and the grating grooves exhibit effects similar to open-ended microwave cavities or waveguide shorts.[4,5] Under these conditions, the group velocity in the grating groove is less than the group velocity in free space; the effect exhibits itself as an effective increase in grating groove depth, resulting in decreased modulation. For a monochromatic source, the interferogram is described by the equation

$$A = 1 + \cos \frac{4\pi d}{\lambda}, \qquad (20\text{-}2)$$

where A is the zero-order illumination, d is the groove depth, and λ the wavelength. If we consider successive maxima in Figure 20-4—an interferogram of

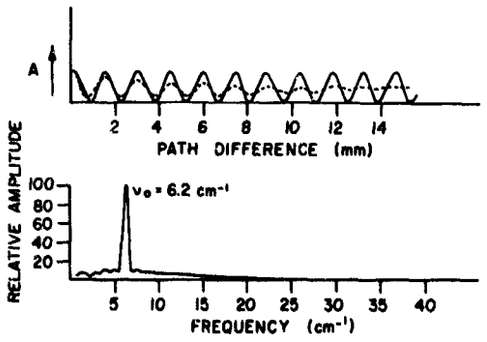


Figure 20-4. Interferogram of Monochromatic Source and Spectrum. The dashed line in the interferogram indicates what could be expected in the presence of "cavity effects"

monochromatic radiation—a detectable change in the magnitude of A, ΔA, should occur if λ changes due to *cavity effects*. From Eq. 20-2,

$$\Delta A = \frac{4\pi d}{\lambda^2} \sin \frac{4\pi d}{\lambda} \Delta\lambda. \qquad (20\text{-}3)$$

For maxima $d = n\frac{\lambda}{2}$ $(n = 0, 1, 2, \ldots)$,

$$\Delta A_n = \frac{2n\pi}{\lambda} \Delta\lambda. \qquad (20\text{-}4)$$

Thus the amplitude of the successive peaks of the interferogram of the monochromatic source should decrease by an amount $\left| 2\pi \frac{\Delta\lambda}{\lambda} \right|$.

We actually tested the grating with the monochromatic radiation of two klystrons operating at 4.5 and 6.2 cm^{-1}. A typical interferogram and the corresponding spectrum are shown in Figure 20-4. The dashed curve in this figure indicates what would have occurred in the presence of cavity effects. But there was no detectable change in the amplitude of the interferogram (solid line) peaks. This was the case for both polarizations (parallel and perpendicular to the long groove direction) and for both klystron frequencies. Tests of this type have been conducted by Vanasse[1] on a similar grating ($a = 1.3$ cm) with a 4.4 mm (2.2 cm^{-1}) klystron. His results indicate a stronger central maximum, indicating the possible presence of cavity effects.

The maximum path difference of our grating is 1.8 cm which results in a theoretical resolving power of 0.27 cm^{-1}.[1,6] The best resolution achieved in practice is between $\Delta\sigma = 0.4$ and $\Delta\sigma = 0.5$ cm^{-1}. From the upper cutoff frequency of $\sigma = 100$ cm^{-1}, the digitizing interval should be no greater than 0.0025 cm. Usually in practice we use the value 0.0019 cm.

The computer program replaces the Fourier integral with a corresponding summation. A triangle apodization function is used. Since we use only the cosine (even) transform, we employ a parabola curve-fitting technique to the central maximum of the interferogram to locate the zero point and to make appropriate phase error corrections.[7] Usually a digitized interferogram contains about 950 data points, each contained in a seven-character word. It takes our time-sharing system 15 minutes to accept this information. Actual calculation takes about 70 seconds, and print-out, including a spectrum plot is usually completed within 30 minutes. It thus takes about 45 minutes to obtain a spectrum from a completed interferogram.

20-3 MEASUREMENTS

We have obtained spectra of atmospheric water vapor absorption, and the emission spectra of an HA3 mercury lamp, both in a vacuum and in our system purged with helium. Also, we have determined the channel spectra of quarts plates with alternating silvered facets deposited on opposite sides.

Figure 20-5 shows the HA3 mercury lamp emission spectra, both in a vacuum and in the dry helium

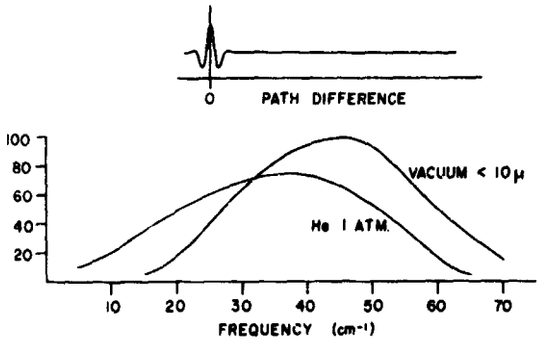


Figure 20-5. HA3 Mercury Lamp Spectrum in Vacuum and Dry Helium System and Typical Interferogram

atmosphere. The spectra appear similar, but the vacuum spectrum has been shifted to the higher frequencies. We attribute this effect to the decreased temperature of the mercury lamp in the second instance, caused by increased convection cooling.

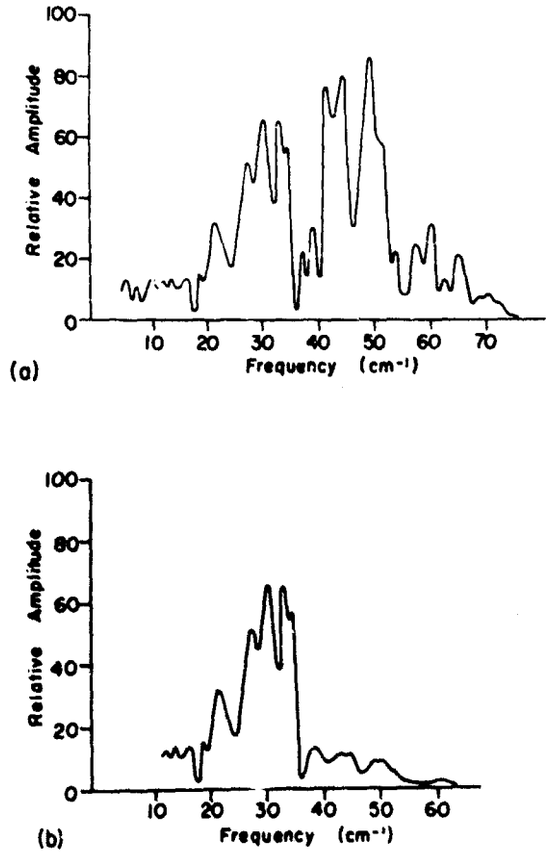


Figure 20-6. (a) Spectrum Obtained of Mercury Arc Radiation Passing Through 3.3 Meters of Laboratory Air. (b) A Yoshinaga Filter is Inserted in the Path of the Radiation. Frequencies higher than 40 cm^{-1} are highly attenuated

Figure 20-6(a) shows HA3 lamp radiation passing through 3.3 meters of atmospheric water vapor. In Figure 20-6(b) a Yoshinaga filter is inserted into the path of the radiation and is seen to cause high attenuation for frequencies higher than 40 cm^{-1}.

Figure 20-7(a) shows how a quartz plate was

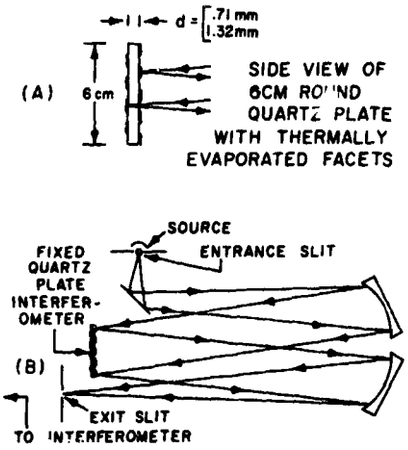


Figure 20-7. Collimator for Quartz Plate Channel Spectrum Measurements

oated by strips of reflecting silver—deposited by thermal evaporation—alternating facets being on opposite sides. Essentially this plate forms a lamellar grating interferometer of fixed geometrical path difference. When radiation at near normal incidence strikes the plate, destructive interference will occur when $k\lambda = 2nd (k = 1, 2, 3 \ldots)$, where λ is the characteristic wavelength, d the thickness of the plate, and n the index of refraction.[8] Thus channel spectra may be obtained which will yield information as to n, since d is invariant, and known; and λ may be calculated from successive minima in the channel spectrum. The collimator for the quartz plate interferogram is shown in Figure 20-7(b). Its entrance and exit apertures are rectangular slits of 0.6 cm width. The grating spacing with this slit width puts higher order spectra outside the exit aperture, so that energy in these orders does not reach the entrance aperture of our interferometer. Two quartz plates, of thickness 1.32 mm and 0.71 mm, were used.

The interferograms obtained with the quartz plates show signatures characteristic of channel spectra. It was determined, optically, that the plates were cut nearly perpendicular to the optical axis. However, from the closely spaced signatures on the interferogram and the modulated cosine structure of the spectra (Figure 20-8), we conclude that our quartz plates exhibited a small amount of birefringence. Such spectra may be represented as the

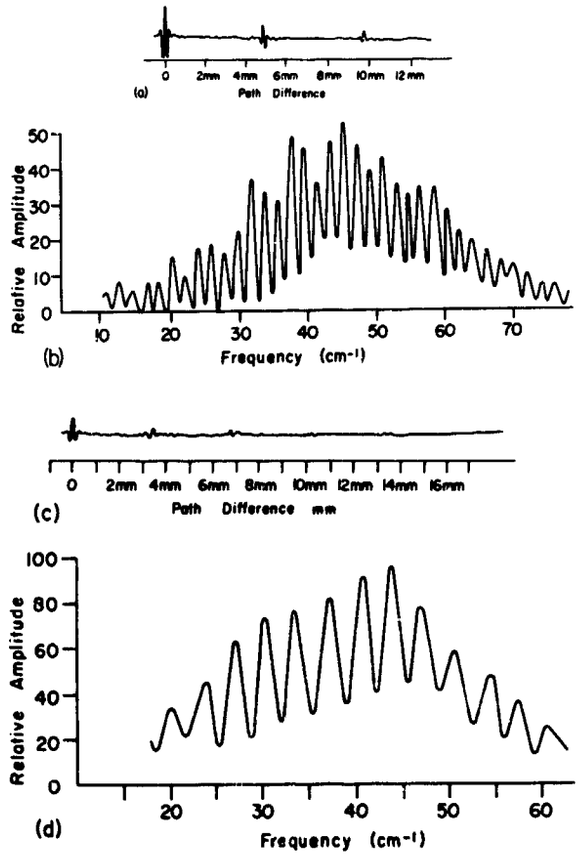


Figure 20-8. (a) Interferogram of 1.32-mm Quarts Plate. (b) Corresponding Spectrum. (c) The Interferogram Obtained With a 0.71-mm Plate Thickness. (d) The Resulting Channel Spectrum

sum of two cosine oscillations of slightly different frequencies according to the trigonometric identity:

$$2\cos[2\pi(n_1+n_2)\sigma d]\cos[2\pi(n_1-n_2)\sigma d]$$
$$=\cos 4\pi n_1\sigma d+\cos 4\pi n_2\sigma d, \quad (20\text{-}5)$$

where n_1 and n_2 are the indices of refraction.[6] From the interferograms and the spectra, we have determined the index of refraction of quartz to be 2.10 for $20<\sigma<60$ cm^{-1}.

Acknowledgments

This work was supported by ONR Contract No. N00014-67-A-0230-0002.

References

1. Vanasse, G. A. (1959) *Interferometric Spectroscopy in the Far Infrared.* Ph.D. Dissertation, The Johns Hopkins University, Baltimore, Maryland.

2. Strong, J., and Vanasse, G. A. (1959) *J. Opt. Soc. Am.* **49** (No. 9): 844.

3. Chaney, L. W. (1967) *Fundamentals of Fourier Transform Spectroscopy,* Technical Report, University of Michigan, College of Engineering.

4. Wolf, E., Ed. (1969) *Progress in Optics,* John Wiley, New York, Chapter 1.

5. Plonsey, R., and Collin, E. (1961) *Principles and Applications of Electromagnetic Fields,* McGraw-Hill, New York, Chapter 10.

6. Loewenstein, E. V. (1960) *Interferometric Determination of Far Infrared Line Width,* Progress Report AFOSR–TN-6065.

7. Tolansky, S. (1948) *Multiple-Beam Interferometry,* Clarendon, Oxford, Chapter 13.

8. Loewenstein, E. V. (1963) *Appl. Opt.* **2** (No. 5): 491.

Discussion

Q. (J. V. Kline, Colorado School of Mines): This morning Dr. Strong said he preferred to turn the lines of the grating horizontal instead of vertical, but he did not say why.

A. (John Strong): When the lamellar grooves are parallel to the plane of incidence, the spherical surfaces being worked "off-axis," the movable facets do not occult the fixed facets, when the moving facets are *out;* but the rim support of the fixed facets occults the moving facets when they are *in.* This gives an asymmetrical interferogram. When the grooves are perpendicular, the preferred orientation, the movable facets shade the fixed facets more and more as they move *out.* Likewise, the fixed facets shade the movable facets as they move *in.* This gives a symmetrical interferogram that is naturally apodized.

Contents

21. An Automated Interferometer Spectrometer: Real-time Data Acquisition and On-line Computation to 0.5 cm^{-1}

Raul Curbelo and Charles Foskett
Digilab, Inc., Cambridge, Massachusetts

21–1 INTRODUCTION

The FTS–14 is an automated Michelson interferometer spectrometer system. By interchanging beam splitters and source detectors, the entire wavelength region from 0.6 μm to 100 μm can be measured. The basic system is designed to operate in the near infrared, from 2.8 μm to 25 μm, at better than 0.5 cm^{-1} resolution. The spectrometer is completely self-contained and is equipped with its own general-purpose data system for control, acquisition, computation, and display.

21–2 HARDWARE CONFIGURATION

21–2.1 Analog Subsystem

The optical head contains a Michelson type of rapid-scanning interferometer. One mirror of the interferometer is mounted on an air bearing and is moved by an electromechanical transducer similar to that of a loudspeaker. The mirror motion is controlled by a constant-velocity servomechanism.

A reference interferometer is mounted behind the moving mirror of the main interferometer. The reference interferometer has two detectors, which look at a laser and a white-light source. The resulting interferograms are used to generate the velocity error for the servomechanism and the sampling information for sampling the main interferometer signal.

The double-beam system has two sample compartments, through either of which the beam passes exclusively. To minimize background effects, the source is modulated by the interferometer before it passes through the sample compartments. The beam is focused at the sample compartment by an off-axis paraboloid.

21–2.2 Digital Subsystem

The signal is digitized by a 15-bit analog/digital (A/D) converter, with a 10-μsec digitization time. The input range of the A/D converter is approximately 10 V peak-to-peak.

Digital data is entered into the data system memory via the central processor, although other alternatives are available. The control processor unit (CPU) and its associated core is the Data General NOVA.

The NOVA is a general-purpose 16-bit computer with a 2.6-μsec cycle time. Typical instructions take 5.6 μsec. There are four fully programmable registers, two of which can be used for indexing. The instruction set is fully complemented, and the system has good basic software such as an assembler, editor, and debugger. The memory can be expanded to 32K words; up to 64 devices can be appended to the input/output (I/O) bus. Data can be entered into memory via a data channel either directly or in addition.

The FTS–14 has two basic configurations: all-core, and disk. The all-core system has 16K of memory, 12K of which is used for data. The disk system has 4K of core memory and 128K of disk memory. Either system can be optionally expanded to 32K of core memory.

Status is determined and control effected through a generalized I/O interface. This interface uses 16 lines for status determination, 16 for digital output control, 8 for inputting alarm conditions to the

spectrometer, and 4 for general condition-sensing. Some of these lines are used by the FTS–14 basic system. The majority remain available for general-purpose laboratory use.

A high-speed multiply/divide unit is used to speed computations. All operations are signed. The basic multiply time is 4.4 μsec. Combined with input and output, a multiply typically takes 20 μsec.

Display is achieved through a Houston DP–1 incremental plotter. The plotter increment is 0.005 in. and plot speed is 300 increments per sec. A second pen is simultaneously run as a wavelength marker.

21–3 SOFTWARE CONFIGURATION

21–3.1 Disk vs All Core

The software configuration is roughly the same in both the all-core version and the disk versions. Advantages of the core system are its compactness and access speed. The advantage of the disk system is its memory size. In each configuration the software is described under three separate headings: (*a*) control and data acquisition, (*b*) computation, (*c*) display.

21–3.2 Control and Data Acquisition

User interaction with the system is implemented through a conversational executive. The conversational executive uses a basic three-letter word system to store system and experimental parameters and to store and execute system commands. An arbitrary number of parameters or commands can be used in the executive.

The status lines sense the condition, or state, of the experiment. Air bearing, air pressure, laser intensity, scan status, beam status, and control-panel switch positions are all monitored by the status lines. Air pressure, laser intensity, and interferometer power are monitored by the alarm lines.

The data-acquisition program sets various states as a function of user-determined system parameters through the control lines. These include the measure/align optics, variable apertures, sample/reference optics, and—optionally—an automatic sample changer. In addition, the scan release/reference condition is controlled through these same lines.

The data are signal-averaged in real time based on a clock derived from the laser reference. The number of scans for which the signal is averaged is a user-controlled parameter.

21–3.3 Computation

The computation portion of the program includes user-controlled off-centered trapezoidal apodization. The FFT program is a mixed-radix (two/four) decimation-in-time version of the Cooley-Tukey algorithm. The coefficients can be computed or tabulated. All arithmetic is binary (fixed-point), with the binary point at the left. Scaling is performed only when required.

The phase correction performed is multiplicative, following L. Mertz,[1] and uses the center portion of the interferogram to obtain smooth reference-phase spectra. The correction is independently performed on both the sample and the reference.

21–3.4 Display

The FTS–14 is a double-beam spectrometer system, readily used in the emission mode. A useful feature of the software configuration is that during data display the data stored in memory are not disturbed. For example, assuming that a reference and a sample spectrum have been computed and stored in memory, there are several plotting options. The user can plot either spectrum as a single-beam spectrum or he can plot transmittance, absorbance, or log absorbance in any sequence, any number of times, without destroying or altering the memory store. In addition, the plots can be scale-expanded, or the data can be smoothed to a lower resolution for a survey-mode spectrum, after which important regions can be replotted at high resolution, still from the same data.

Either beam, from front (F) or back (B), can be used as the sample beam, and so F/B or B/F can be plotted. In addition, a reference can be stored, and so F/F_0 or B/B_0 are also acceptable modes.

21–4 SUMMARY

The FTS–14 marks a new turn in Fourier spectroscopy. A small general-purpose computer and an interactive software configuration provide digital control over resolution, signal averaging periods, sample/ratio modes, mirror scan and retrace, apodization, and display mode. Most important, the spectrum is fully transformed and phase-corrected in seconds.

References

1. Mertz, L. (1965) *Transformations in Optics*, Wiley.

Discussion

Q. (Lawrence P. Giver, NASA Ames Research Center): What detectors are provided for this instrument?

A. (S. T. Dunn, Digilab, Inc.): Most commercially available detectors are easily interfaced to either the dual-beam sampling system or directly with the spectrometer. As you saw from the slides, the interfacing of cooled detectors is straightforward. We have provided the following detectors with various systems:

1) TGS
2) Thermistor Bolometer
3) Hg Cd Te (77° K)
4) PbSe (cooled and at room temperature)
5) PbS (cooled and at room temperature)
6) InSb (cooled and at room temperature)
7) Silicon cell
8) PM tube
9) Ge cell
10) Cu- and Au-doped Ge

Q. (James R. Aronson, Arthur D. Little, Inc.): Do you have any trouble with the TGS detector exceeding its Curie point?

A. (S. T. Dunn): The Curie point is about 45° C and we have never exceeded this. Thus, we have had no problems with the TGS in this area. In many cases a failure of the TGS can be corrected by repotting the detector. We anticipate that the TGS detector will be very reliable in the field.

Q. (Harry Andrews, U. Southern California): How many data points do you use per spectrum? Does your hardware use fixed or floating point?

A. (S. T. Dunn): For the high resolution spectra we record 16,000 points and calculate 16,000 output points, then interpolate these to 64,000 points for plotting. The basic data system can perform a 64,000-point transform. With additional disks and more double-precision calculations, we can expand this to much larger transforms. The calculation times for the data system of less than 30 sec for an 8K transform compare very favorably with the chart of computation times for large computers presented by Dr. J. Connes yesterday. The arithmetic contains some floating-point calculations for precision during critical operations.

Contents

22. Beam Splitter Optimization for Fourier Spectroscopy

A. G. Tescher
General Dynamics
Pomona, California

Abstract

In an analytic study of properties and limitations of beam splitter designs for the infrared part of the spectrum (1.2 μm through 30 μm), simple designs are compared with multilayer configurations. Polarization effects are found significant in all cases.

22–1 INTRODUCTION

The basic advantage of Fourier spectroscopy is that the number of simultaneously observable resolution elements and their corresponding efficiencies are independent—that is to say, the multiplex principle applies. The efficiency for any particular portion of the spectrum may be defined here as the fraction of the available radiation utilized by the detector.

To achieve high efficiency over a wide spectral band, three instrumentation aspects are important for practical Fourier spectroscopy: (*a*) sufficient optical alignment must be provided over the entire spectral region of interest, (*b*) the appropriate photoelectric detector must be available, and (*c*) the spectral efficiency of the required beam splitter must be high over as wide a band as possible. Also involved are indirect problems such as the required dynamic range of the recording system, which increases proportionally with the width of the observed spectral band. Most of these problems have been discussed analytically in the literature[1] for application in the infrared portion of the spectrum. The dynamic range requirements and solutions thereto are discussed in Chapter 27 of this collection (Fourier Spectroscopy Applied to Field Measurements, by G. W. Ashley and A. G. Tescher).

Surprisingly, one basic element in all Fourier transform spectrometers—the beam splitter—has received relatively little attention. Although a simple problem for narrowband application, the beam splitter configuration for wideband operation is not well established. For the purpose of this discussion, we define as wide- and narrowband the cases where the bandlimit ratio exceeds 2 or is less than 2, respectively. A common wideband application for the near-infrared, for example, is the region from 1.25 μm to 5 μm since the bandlimit ratio approaches 4.

The basic objective of this study was to determine simple beam splitter configurations for very wide band applications, that is, for cases where the bandlimit ratio approaches 10. In addition to specifically considering the 1.2-μm to 12-μm and the 3-μm to 30-μm bands, some effort was made to study the effect of the beam splitter on the polarization of the radiation that passes through it.

22–2 DESIGN APPROACH

The typical beam splitter (Figure 22-1) consists of a transparent substrate, usually a low-index material, on which there has been deposited a thin film coating consisting of a single layer or numerous elements. A compensating plate, usually of the same material as

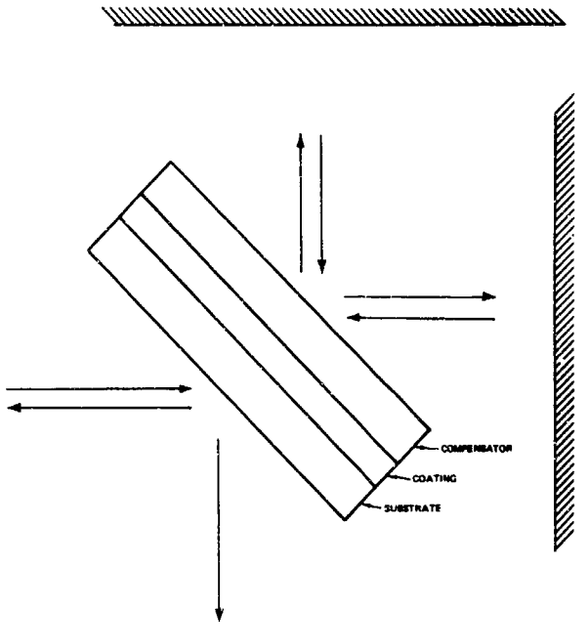


Figure 22-1. Schematic Representation of Beam Splitter Arrangement

the substrate on which the thin film was deposited, completes the beam splitter.

In the conventional Michelson type of configurations, the beam undergoes one reflection by and one transmission through the beam splitter. Consequently, the energy available at the photoelectric detector is proportional to the product of power transmission and reflection. This product has a maximum value of 0.25, which defines the design goal for the optical efficiency of the beam splitter; both reflection and transmission coefficients should therefore be as close as possible to 50 percent at all points of the useful spectral region of the beam splitter.

To meet this objective, we first considered only dielectric configurations. Other materials are obviously undesirable because of energy losses due to absorption. In addition, it can be shown that for dielectric coatings the complementary interferograms are identical and 180° out of phase, which is in fact the optimum arrangement (Mertz,[1] page 34). [See Appendix A for a proof for the 180° phase shift.]

For the detail design, we have developed a rather complex computer program that allows us to calculate the optical characteristics of various thin film configurations. Only a very brief qualitative description will be given here. As the basis of the algorithm, we use the technique of characteristic matrices.[2,3] Accordingly, any thin film layer can be described exactly by a 2×2 complex matrix, and the composite structure is equivalent to the product of the individual matrices. The final transmission and reflection parameters can then be expressed in terms of the matrix elements of the composite structure. The individual matrix elements are given in closed form for homogeneous layers. When the refractive index itself is a continuous variable of position, the matrix elements are calculated by numerical integration. Since the efficiency of the beam splitter is given by the product of power reflection and transmission, the described program lends itself easily to the optimized design of beam splitter geometries.

It should be emphasized that the calculations thus performed yield exact optical characteristics within the validity of the macroscopic theory. If we can describe each particular component by its macroscopic properties, that is, by its index of refraction, all boundary conditions are automatically satisfied.

The computer program handles either polarization, any number of layers, dielectric, or absorbing components, and arbitrary incident angles.

22–3 RESULTS

22–3.1 Simple Configurations for the 1.5-μm to 5-μm Region

Figure 22-2 indicates the optical characteristics of

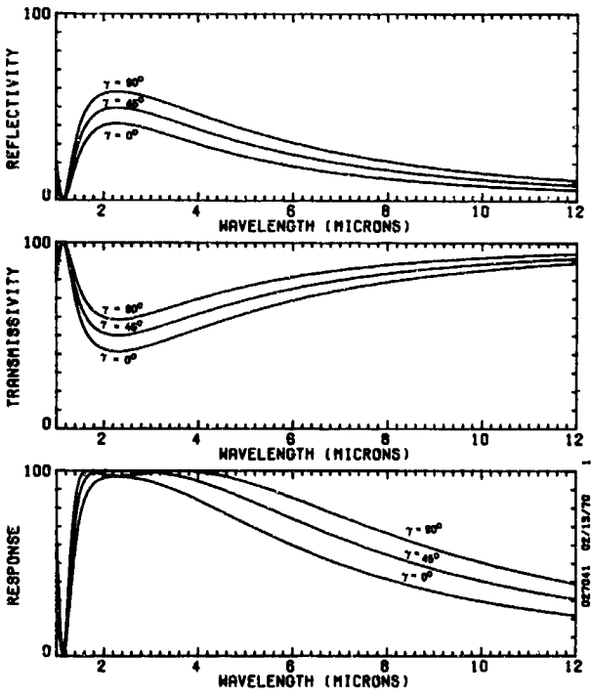


Figure 22-2. Optical Properties of Calcium Fluoride/Silicon/Calcium Fluoride Beam Splitter

a conventional beam splitter consisting of calcium fluoride coated with a 0.17-μm-thick silicon layer; a calcium fluoride compensating plate is assumed in direct contact with the coating. Here, as well as in further examples, the incident radiation is at 45°. The

refractive indices for calcium fluoride and silicon were taken to be 1.42 and 3.44, respectively. The upper two graphs show the power reflection and transmission through the beam splitter. The bottom graph is the effective response, which is calculated as the normalized product of the transmission and reflection curves, that is, 100 percent corresponds to the theoretical maximum of 0.25. The calculation is performed for three polarization angles: 0°, 45°, and 90°. The 0° and 90° correspond to TM and TE polarizations, respectively. The 45° polarization indicates an equal amount of power in either polarization. The efficiency is calculated from the product of the appropriately weighted power reflection and transmission curves.

If we accept 80 percent efficiency as a design goal, this simple beam splitter configuration is very satisfactory from 1.5 μm to 5 μm. For the longer wavelengths, the relative efficiency seems to be considerably affected by the type of polarization chosen. We have built this configuration and find that the measured reflection and transmission characteristics correlate extremely well with those calculated.

22-3.2 Simple Configuration for the 4-μm to 16-μm Region

The materials chosen for the 4-μm to 16-μm range were sodium chloride for the substrate as well as for the compensating plate, both in contact with a 0.4-μm-thick coating of germanium. The corresponding refractive indices are 1.5 and 4.0, respectively. The characteristic features of the response (Figure 22-3) show a null point at 3.2 μm, and a monotonic decrease at wavelengths longer than 12 μm. In addition, at the longer wavelengths the curves separate significantly with polarization. We have built this configuration also, and find here also that the experimentally measured values correlate well with those calculated.

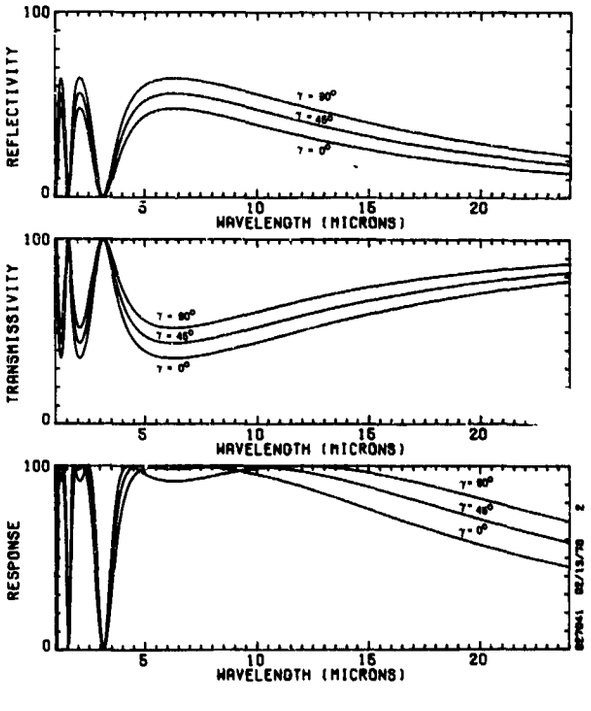


Figure 22-3. Optical Properties of Sodium Chloride/Germanium/Sodium Chloride Beam Splitter

22-3.3 Beam Splitter for the 1.5-μm to 12-μm Region

The important question arises whether, by minor modification of the configuration, the simple beam splitter can be utilized over a wider spectral band. The answer is that even with the addition of a single thin film layer, considerable improvement can be achieved. Figure 22-4 shows this for the calcium fluoride/silicon beam splitter. The compensating plate is separated from the 0.17-μm coating by a distance of 2.5 μm, that is, an air gap is used as a second, low-index layer. The obvious effect is the induced modulation in the reflection and transmission curves. The response curves show somewhat more complex behavior than was observed for the simpler configuration. We may, however, observe considerable improvement in the spectral efficiency at longer wavelengths. For random polarization the spectral efficiency is still above 50 percent at 12 μm. Consequently, this beam splitter configuration is useful over the range of 1.5 μm to 12 μm. We should also note the strong variation in the efficiency curves as functions of polarization. In the case of 90° polarization, the average value of the response curve is above 80 percent in the 1.5-μm to 12-μm region.

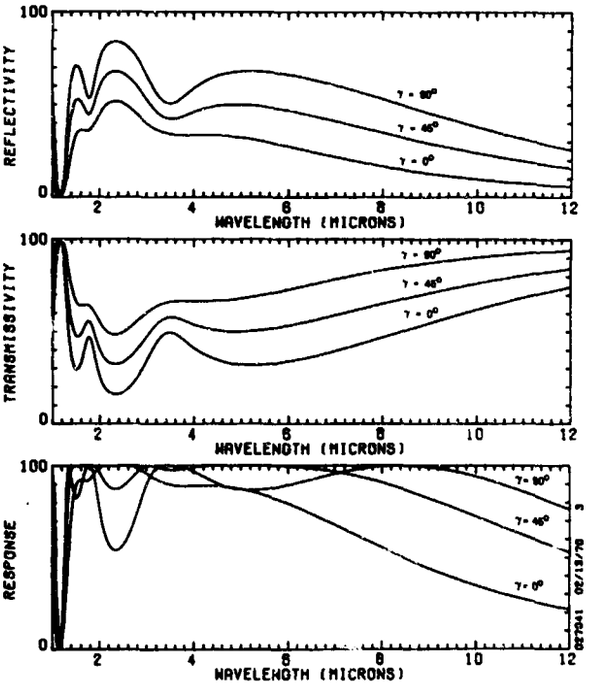


Figure 22-4. Optical Properties of Calcium Fluoride/Silicon/Calcium Fluoride Beam Splitter With 2.5-μm Air Gap

22-3.4 Beam Splitter for the 3-μm to 30-μm Region

For comparison with the simple design for the wideband configuration, the optical properties of a beam splitter utilizing multilayer coatings were analyzed. The example discussed here is the one developed by P. L. Heinrich et al.[4] for application in the region from 3 to 30 μm.

Thirteen layers, of three different materials (Ge, TlBr, KRS-5), were used. The optical properties are shown in Figure 22-5. The overall thickness was over

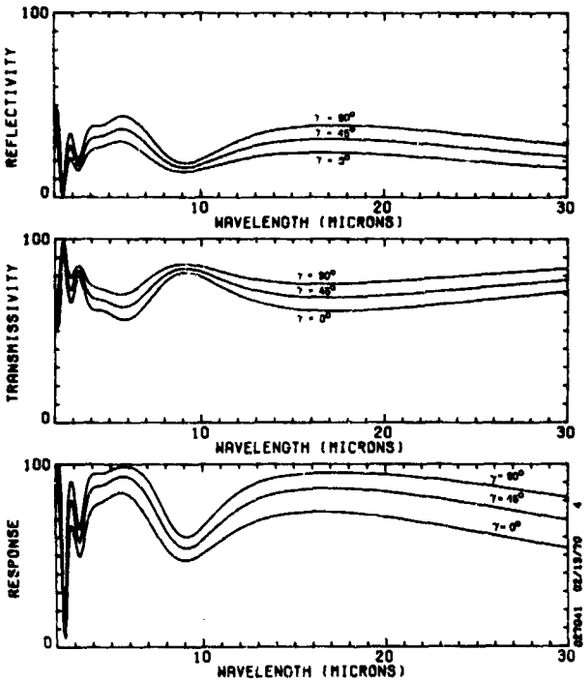


Figure 22-5. Optical Properties of Multilayer Beam Splitter

1.7 μm, which is considerable. In fact, the discussed configuration was a limiting case in terms of coating technology. Systems consisting of more than 13 layers became mechanically unstable. The calculated optical properties are good, but do not really exceed the previously discussed simple configuration with the air gap. The optical properties vary as a function of polarization but somewhat less than for the simple design. The response characteristic is good at long wavelengths; in the 10-μm region, however, a considerable drop (to 50 percent) can be noticed. The effective response is marginal at wavelengths shorter than 4 μm. Consequently, the ratio of the passband limit is only 30:4, or 7.5, even if we ignore the deficiency in the 10-μm region. The similar passband limit ratio for the simple design with an air gap (Figure 22-4) is 12:1.4, or approximately 8.6.

22-3.5 Parametric Effect of the Air Gap

We have indicated that the introduction of a second element into the simplest beam splitter configuration can very favorably alter its optical properties. In Figure 22-6 we present the effect of the width of the

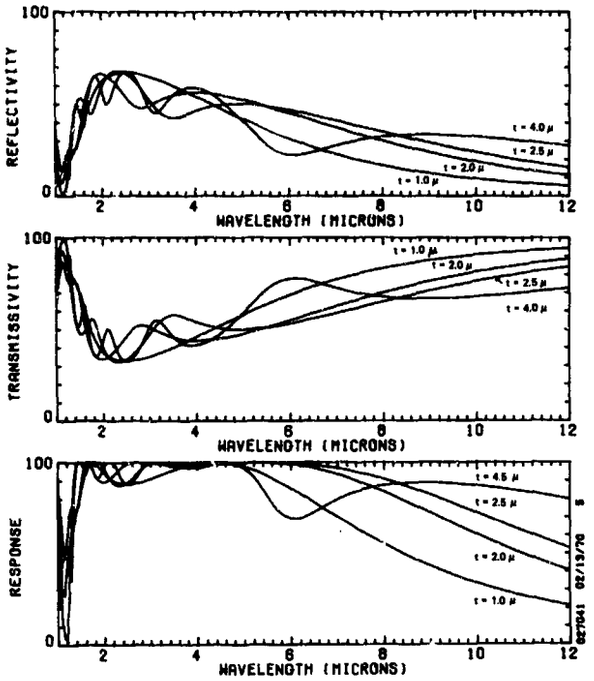


Figure 22-6. Optical Properties of Calcium Fluoride/Silicon/Calcium Fluoride Beam Splitter for Air Gaps of Four Different Widths

air gap on the characteristics of the beam splitter. The substrate and compensating plates are calcium fluoride, and the coating is a 0.17-μm-thick silicon layer. The polarization is 45°, corresponding to random polarization. The air-gap thicknesses considered are 1.0, 2.0, 2.5, and 4.0 μm.

We notice significant variations in optical characteristics. Consequently, the air-gap width can be used to optimize the beam splitter for specific spectral regions of interest. This is an important practical consideration since operation over a very wide spectral band may be required only occasionally, with the majority of investigations confined to a narrower optical region. The approach in this case would be to achieve near 100 percent efficiency in the narrower spectral region while maintaining at least a marginal efficiency over the wider band. The additional degree of freedom does allow this design consideration.

22-4 CONCLUSIONS

The analytic and experimental evidence of this work has shown that simple, single-element beam

splitters can be fabricated with high efficiency for operation in the bandpass from 1.5 to 5 μm and from 4 to 16 μm. The addition of only a single, low-index film element such as an air gap is shown to be sufficient to produce a very satisfactory beam splitter configuration for operation over a band of 1.5 to 12 μm or wider. The performance characteristics of such a relatively simple configuration are quite compatible with those of far more complex systems using thin film stacks of more than 10 layers.

References

1. Mertz, L. (1965) *Transformations in Optics*, Wiley.

2. Born, M., and Wolf, E. (1959) *Principles of Optics*, Pergamon Press, Sec. 1.6.

3. Tescher, A. G. (1966) Computer techniques for the analysis of inhomogeneous thin film configurations, *JOSA* **56**: 554.

4. Heinrich, P. L., et al. (1967) *Development of an All Dielectric Infrared Beamsplitter Operating in the 5 to 30 Micron Region*, NASA CR-703-1967.

Appendix 22-A

We want to show that the phase difference at a simple dielectric layer is 90° between the reflected and transmitted plane waves (see Figure 22-A1).

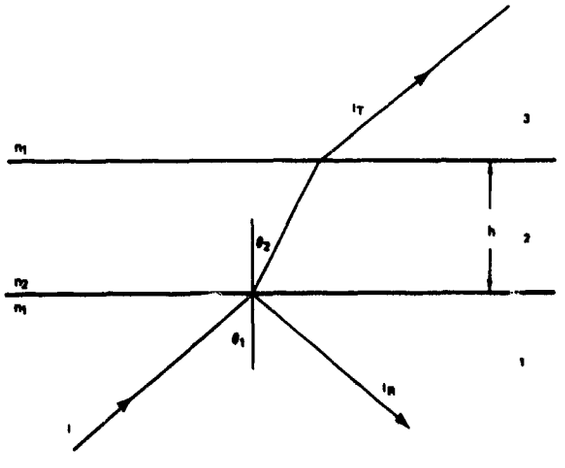


Figure 22-A1. Propagation of an Electromagnetic Wave Through a Dielectric Layer

Proof. The phase changes upon reflection and transmission are given by:[2]

$$\tan \delta_r = \frac{-r(1-r^2)\sin 2\beta}{r(1+r^2)-r(1+r^2)\cos 2\beta};$$

$$\tan \delta_t = \frac{1-r^2}{1+r^2}\tan \beta$$

where

$$r = \frac{n_1 \cos \theta_1 - n_2 \cos \theta_2}{n_1 \cos \theta_1 + n_2 \cos \theta_2}$$

and

$$\beta = \frac{2\pi}{\lambda} n_2 h \cos \theta_2.$$

By straightforward trigonometric manipulation

$$\tan \delta_r \tan \delta_t = \frac{-\sin 2\beta \tan \beta}{1-\cos 2\beta} = -1.$$

Therefore,

$$\tan (\delta_r - \delta_t) = \frac{\tan \delta_r - \tan \delta_t}{1+\tan \delta_r \tan \delta_t} = \infty$$

and

$$\delta_r - \delta_t = 90°.$$

Consequently, the complementary interferograms are 180° out of phase.

Discussion

Q. (Paul D. Feldman, Johns Hopkins University): How thick must the substrate be in order to obtain optical flatness and how much absorption does this introduce?

A. (A. Tescher): Our substrates were 0.4 in. thick. They should be so chosen as not to produce any absorption.

Q. (Paul D. Feldman): What about using Mylar as a substrate?

A. (A. Tescher): We have only considered the 1- to 30-μm region. It is unlikely that the Mylar could provide sufficient optical flatness.

Comment (Hugh A. Gebbie, National Bureau of Standards): Films can be evaporated on Mylar substrates. We can then evaporate layers on both sides and exploit the low index of the Mylar in the same way that you have used an air film.

Comment (A. Tescher): The suggestion may be appropriate for the far infrared; however, it is unlikely that we could maintain the required optical flatness for shorter wavelengths.

Q. (Dr. Alain L. Fymat, Jet Propulsion Laboratory): You have described for us the optical characteristics of various beam splitters for different states of polarization of the incident radiation. However, since the incidence of the radiation on the beam splitter is not along the normal, a residual polarization will therefore also be introduced. I was wondering whether your computer program has been written to handle also this polarization.

A. (A. Tescher): The incident radiation was assumed at 45° (the computer program can handle any incident direction), and the indicated efficiency curves refer to the appropriate polarizations.

Q. Have the air-gap CaF_2 beam splitters been fabricated?

A. (A. Tescher): No; however, no difficulty would be anticipated in building one.

Contents

23. The Nimbus III Michelson Interferometer

R. A. Hanel and B. Schlachman
Goddard Space Flight Center

F. D. Clark, C. H. Prokesh*, J. B. Taylor†, and W. M. Wilson
Texas Instruments, Inc.

L. Chaney
University of Michigan

Abstract

The Michelson interferometer flown on Nimbus III in April 1969 has obtained infrared emission spectra of the earth and its atmosphere within 400 and 2000 cm^{-1} (5 and 25 μ). Spectra of good quality have been recorded with a spectral resolution corresponding to 5 cm^{-1}. This paper contains a discussion of the design of the instrument including the optical layout, the phase-locked loop operation of the Michelson motor, and the functioning of the reference interferometer. The methods of data reduction and in-flight calibration are demonstrated on sample spectra recorded while in orbit around the earth.

23-1 INTRODUCTION

The use of a Michelson interferometer as a spectrometer is relatively recent. Michelson's original technique of observing the "visibility of fringes" gives spectral information in certain special cases only. It was not until the 1950's that interest was revived in the Michelson interferometer as a spectral analyzer. After Fellgett[1] had shown the multiplex advantage and Jacquinot and Dufour[2] had called attention to the large-area solid-angle product of the interferometer, Fourier spectroscopy was advanced by several groups, notably in England, France, and the United States. The state-of-the-art at that time is reflected in papers by Jacquinot[3] and Mme. Connes.[4] The paper by Mme. Connes was of great help in the design of the instrument being discussed. Interested readers may find more recent surveys on Fourier spectroscopy by Mertz,[5] Lowenstein,[6] and Vanasse and Sakai.[7]

The rapid progress that satellite technology experienced in the late 1950's and early 1960's stim-

* Now with Walston and Co.
† Now with the Space Science Department, Rice University.

ulated interest in remote sensing of atmospheric parameters. King[8] and Kaplan[9] suggested methods of determining atmospheric temperature profiles from satellites. For that purpose mathematical techniques and tools were developed in the following years, notably by Wark and his co-workers (see, for example, Wark and Fleming[10] and references therein, King,[11] Chahine,[12] and Conrath[13]). Besides the 15-μ band of CO_2, which is well suited for the temperature analysis, the spectral intervals of water vapor (5 to 9 μ and beyond 17 μ) and ozone absorption (9 to 10 μ) and other spectral regions of low absorption are also of importance to atmospheric research. A careful analysis shows that most parts of the thermal emission spectrum provide useful information. Therefore, the recording of a large portion of the emission spectrum seemed desirable, and it was natural to apply the instrumental advantages of the Michelson interferometer to atmospheric research from space.

The Infrared Interferometer Spectrometer (IRIS) flown on Nimbus III was not the first interferometer which has seen operation in space. Several manned and unmanned satellites carried Michelson interferometers,[14,15,16] but so far only a few spectra have been published from these flights, a fact which emphasizes the difficulty of the measurement.

At the beginning of the project discussed here, it was not clear that a wide spectral range (approximately 5 to 25 μ), a relatively high spectral resolution (approximately 5 cm^{-1}) and a high precision (approximately 1 percent) could be combined into a small instrument, which must also be capable of withstanding the launch and space environment. None of the then available instruments came close to meeting these requirements, and knowledgeable people doubted that such an instrument could be built. At that time groups from the University of Michigan and Goddard Space Flight Center (GSFC) combined forces to prove the feasibility of the task. An instrument, IRIS A, was jointly built and flown on a high altitude balloon in 1966. Encouraging results were reported by Chaney et al.[17] and Hanel and Chaney.[18] Thereafter, the Michigan group continued to develop IRIS A and balloon-borne techniques, while GSFC with the help of Texas Instruments, Inc., developed the IRIS B for space flight.

The purpose of this paper is to describe the design, calibration, and performance of the IRIS B instrument. The first flight model of the instrument was lost when Nimbus B failed to achieve orbit in March 1968. The spare flight model was launched in April 1969 on Nimbus III and produced data for about 3½ months, until the thermistor bolometer failed. Preliminary results have been reported by Hanel and Conrath.[19] Further analysis of the data will be published elsewhere.

23-2 PRINCIPLES OF OPERATION

The essential part of the Michelson interferometer, as shown in the simplified diagram in Figure 23-1,

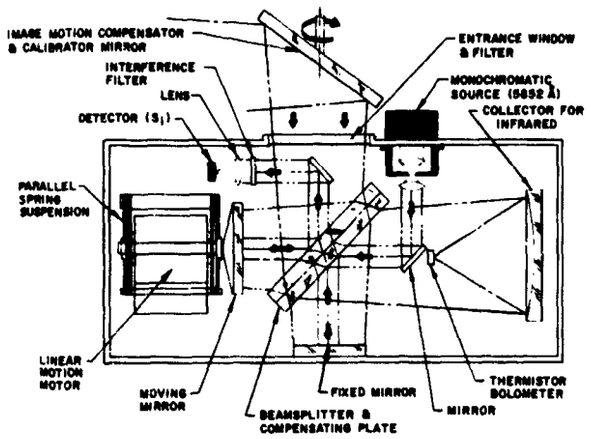


Figure 23-1. Simplified Diagram of IRIS. The image motion compensation and calibration mirror can be oriented so that IRIS sees earth, deep space, or an on-board blackbody

is the beamsplitter, which divides the incoming radiation into two approximately equal components. After reflections on the fixed and moving mirrors, respectively, the two beams interfere with each other with a phase difference proportional to the optical path difference between both incident beams. The recombined components are then focused onto the detector, where the intensity is recorded as a function of path difference δ. For polychromatic radiation and neglecting constant terms, the signal at the detector, called the interferogram, is

$$i(\delta) = \int r_\nu (B_\nu - B_{\nu,i}) \cos\,(2\pi\nu\delta - \phi_\nu)\, d\nu. \quad (23\text{-}1)$$

The recorded amplitude $i(\delta)$ is proportional to the responsivity, r_ν, and the difference in radiance between the scene B_ν, within the field of view and the Planck function $B_{\nu,i}$, corresponding to the instrument temperature T_i. The phase ϕ_ν is slightly wavenumber dependent. Typical interferograms recorded by IRIS in orbit are shown in Figure 23-2. Interferograms 1, 2, and 4 were taken while viewing the earth; interferogram 3 was taken while viewing a warm blackbody. Interferograms 1 and 4 are arctic cases and interferogram 2 is a hot desert case. Each interferogram is digitized and temporarily stored in the spacecraft for transmission to the ground station. The spectrum B_ν is reconstructed in a ground-based computer by applying the inverse Fourier transformation and the calibration procedure to be discussed.

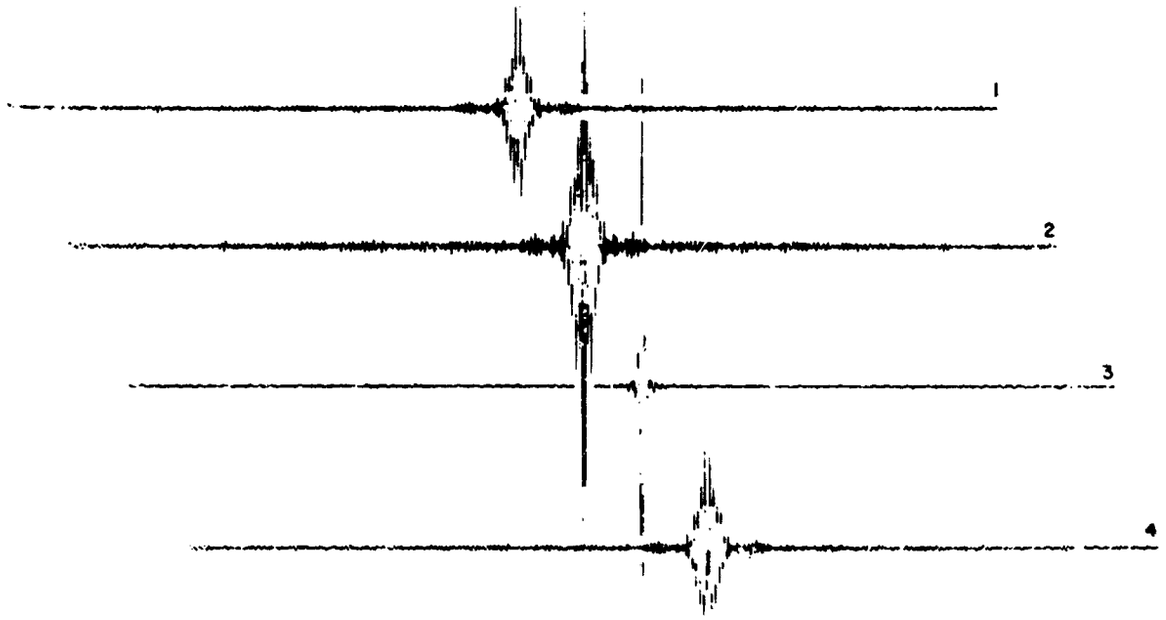


Figure 23-2. Typical Interferograms Recorded by IRIS in Orbit

23-3 DESIGN CONCEPT AND REQUIREMENTS

The functioning of the Michelson interferometer is straightforward in principle, but rather careful design and precise manufacturing techniques are necessary if all advantages of the interferometer are to be realised. The design was dictated partially by the performance requirements and partially by the space environment and spacecraft interface. The desire to resolve five wavenumber intervals determined the distance, 0.2 cm, traveled by the moving mirror. The spectral band to be covered influenced the choice of the beamsplitter substrate, potassium bromide (KBr), and of the detector, a thermistor bolometer. The required spectral resolution and the highest wavenumber of interest determined the field of view; a 4° half cone angle. The desire to operate the instrument at the midrange of expected brightness temperatures, in order to conserve dynamic range, dictated an operating temperature of about 250° K.

The required precision demanded a long integration time per measurement as well as a wide dynamic range in the data transmission process. Analog transmission of the data was ruled out by the dynamic range requirement of about 2000:1 and only a quantization procedure with two scales was considered adequate. The relatively long integration time of 10 sec, combined with the orbital velocity of the satellite, required image motion compensation.

An on-board calibration procedure was considered mandatory to achieve the desired accuracy of about 1 percent of the measured radiance. As will be shown later, without this calibration procedure the experiment would have been a failure. A 1-percent accuracy in the measurement of radiances is difficult to achieve in the laboratory. To approach this accuracy in a remotely operated instrument in space was probably the most ambitious part of the program.

The importance of a linear mirror motion was recognized in the beginning. Sufficiently constant velocity of the Michelson mirror requires feedback control, which was implemented using a velocity transducer and a reference or fringe control interferometer. The signal from the reference interferometer served also as a sample initiator for the analog-to-digital conversion process. Moreover, synchronization of the data stream with the spacecraft clock could be accomplished by a phase locked loop, which slaves the mirror motion to the highly stable clock frequency.

Weight, power, form factor, and the environmental test procedures commonly imposed on space flight instruments, placed further constraints on the design. The weight of the IRIS B is 14.5 kg including the electronic module, and 16 W are necessary for operation. The form factor was determined, in part, by the temperature control requirements. As a result IRIS was mounted under the spacecraft where it could have an unobstructed view of space (thermal heat sink parallel to pitch axis), earth (parallel to yaw axis), and space (calibration point parallel to roll axis). Although IRIS was optimized to operate in

flight at 250° K, practical considerations of qualification testing required that it operate at room temperature also. Furthermore, the interferometer had to survive sinusoidal and random vibration in all three axes without loss of alignment.

The instrument to be constructed had to consist of the following elements, which will be discussed in more detail below:

(1) Optical module including beamsplitter, and fixed and moving mirror,

(2) Moving mirror drive assembly with feedback control,

(3) Reference interferometer and phase-locked loop,

(4) Infrared detector and main data channel,

(5) Logic and control system,

(6) Calibration and image motion compensation elements,

(7) Power system,

(8) Thermal control.

The block diagram in Figure 23-3 shows the interaction of the individual elements of the system.

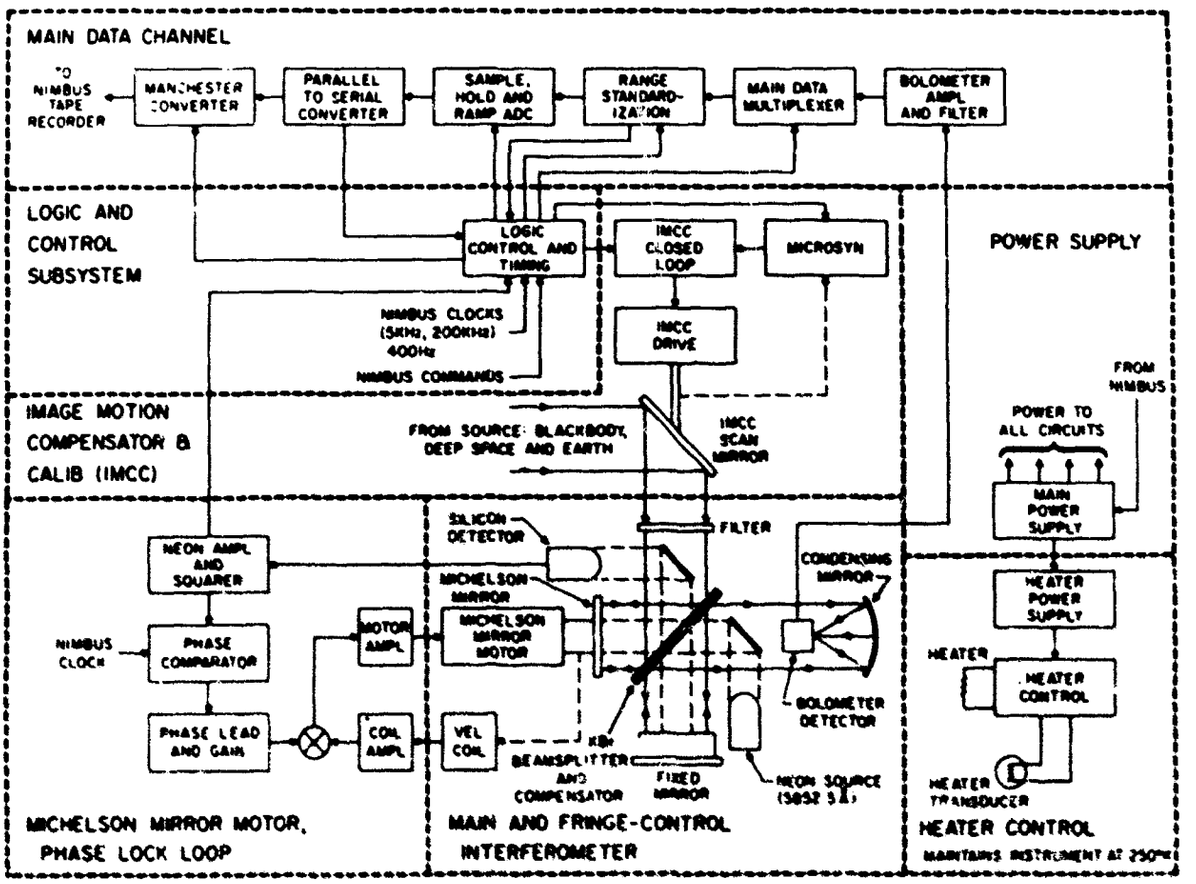


Figure 23-3. Block Diagram of IRIS. The motion of the Michelson motor is phase-locked to the Nimbus clock

23-4 INSTRUMENT DESIGN

23-4.1 Optical Module

The optical module consists of two main components referred to as the Michelson-mirror housing and the fixed-mirror housing. The Michelson-mirror drive, the entrance window, and the detector of the reference interferometer are all mounted on the first housing, shown in the foreground of Figure 23-4. The fixed-mirror housing contains the nonmoving mirror of the interferometer, the condensing mirror, the thermistor bolometer, and the reference interferometer source, all visible on the right side of Figure 23-4. Both housings were machined from 6061-T6 aluminum with walls made as thick as possible to provide rigidity, a large thermal time constant, and small temperature gradients within the instrument.

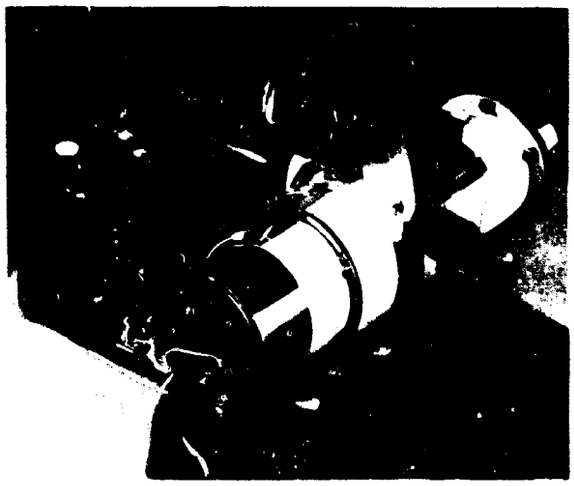


Figure 23-4. Infrared Interferometer Spectrometer (IRIS) Flown on Nimbus III. The upward-looking port is for viewing earth, and the other port is for viewing deep space

The alignment philosophy developed during the design of IRIS depends primarily on the preassembly tolerances and the dimensional stability of the housings, with some fine alignment accomplished by adjustment of the fixed mirror. All mating surfaces that affect alignment were hand lapped to optical tolerances. A pair of wedged spacers allowed fine corrections of the fixed mirror up to three minutes of arc. IRIS was aligned in a walk-in cold chamber at the operating temperature of 250° K. Then the unit was cycled between 300° K and 250° K several times, to uncover any alignment hysteresis.

The Michelson-mirror housing is recessed at the joining flange to receive the beamsplitter mount, which has an optically flat reference surface. An O ring located on the other side loads the beamsplitter mount against an equally well-polished surface of the fixed-mirror housing. The beamsplitter assembly consists of two 0.95-cm thick potassium bromide (KBr) substrates. One substrate carries the multilayer infrared beamsplitter coating, with a separate coating for the fringe control interferometer located in the center. The other substrate serves as a compensator. Transparent films protect both substrates against damage from moisture.

The beamsplitter coatings for the infrared consist of four main layers, each one a quarter wave thick at 5 μ. The refractive index of each layer was chosen to increase from the index of the substrate ($n = 1.53$) to the index of the last layer ($n = 4$). Layers two and four could be made of available materials, KRS 5 ($n = 2.47$) and germanium ($n = 4$), respectively. Layers one and three required indices of 1.94 and 3.15, for which suitable materials do not exist. These layers were made of six sublayers each. The sublayers were $\frac{1}{4}$ of a wavelength thick and involved layers of lead fluoride (PbF_2) and thallium bromide (TlBr) for main layer one, and TlBr and germanium (Ge) for main layer three. The actual beamsplitter surface is the outermost surface of the assembly. All the other layers may be considered antireflection coatings to prevent reflections from the inner surface of the germanium layer. The beamsplitter was designed and fabricated by Perkin-Elmer.

All mirrors of the interferometer are gold coated. The fixed interferometer mirror is the entrance pupil for the optical system. It has an effective circular aperture of 3.5-cm diameter. An ellipsoidal mirror collects the energy from the interferometer and focuses it onto the infrared detector, a thermistor bolometer which serves as the exit pupil.

23-4.2 Moving Mirror Drive

The Michelson mirror is attached to one end of a shaft moving linearly in the direction of its main axis of symmetry. This linear dc motor has a stationary magnetic circuit around a moving drive coil. A velocity transducer, in the form of another axially moving coil in a second magnetic gap, is attached to the other end of the drive shaft. The whole assembly is supported on parallel springs, which ensure that the mirror moves with minimum tilt and without bearing friction.

23-4.3 Reference Interferometer and Phase-Locked Loop

Coaxial with the infrared interferometer is the reference interferometer, which provides velocity control and pulses for sampling the interferogram. Both interferometers share the fixed-mirror, moving-mirror, and beamsplitter substrate. The light source of the reference interferometer is a small neon discharge lamp. Lenses collimate the radiation from the source, and 45° mirrors reflect it into and out of the interferometer path. An interference filter isolates the desired line (0.58525 μ) and another small lens focuses the beam onto a silicon photovoltaic cell. This particular line of neon was convenient, since it was found to be sufficiently strong and narrow, and it permitted visual alignment of the interferometer.

A command from the logic and control system sets the Michelson mirror into linear motion, which generates a sine wave at the output of the silicon detector and a direct current signal at the velocity transducer. The reference signal (625 Hz) is divided in frequency by two, and the positive zero crossing initiates the sample and hold command to the data channel. The reference frequency is also compared in phase with a spacecraft-derived clock frequency of 625 Hz. A signal proportional to the phase error is passed through a gain-adjust network to an integrator and then to a summing point at the current amplifier for the drive coil. The signal from the velocity-control transducer feeds into the same point. The closed-loop operation also reduces the sensitivity of the interferometer to external vibration, a very important point.

The servo system and the optical-thermal design, which maintained alignment over a 50° C temperature

range, were the most difficult individual development tasks of the program. Undoubtedly the phase-lock loop control of the moving mirror contributed considerably to the quality of the spectra obtained.

23-4.4 Infrared Detector and Main Data Channel

A 500-V supply biases the thermistor bolometer at the operating temperature of 250° K. At room temperature, where the instrument must also operate, the bias reduces to 125 V to account for the lower bolometer impedance at that temperature. The bolometer is mounted on spiders in front of the collecting mirror. The frequency response of the high-impedance, low-noise preamplifier mounted behind the collector compensates for the high frequency roll-off of the bolometer. The combined response of the bolometer and amplifier is essentially flat in the frequency range of interest (18 to 73 Hz). A six-pole, active bandpass filter follows the bolometer amplifier and removes noise outside the bandwidth of the data channel. At the lowest aliasing frequency (239.5 Hz) the filter attenuation is about 20 dB, with a roll-off of 18 dB per octave.

The interferogram is multiplexed with auxiliary data, such as the temperature of the calibration blackbody, to provide required information for data reduction within the same frame. The dynamic range of the data channel dictated the use of a gain-changing circuit before the analog-to-digital converter. A dynamic range of 2000:1 is required to process the signal without adding a significant amount of quantization noise. The range standardization circuit has a high-gain and a low-gain region. In each data word one bit, called "gain bit," identifies which one of the regions was active. This gain-changing circuit in effect increases the dynamic range of the analog-to-digital converter by a factor of 10 without a serious loss in signal-to-noise ratio in the computed spectrum.

The analog-to-digital converter is an eight-bit ramp type design preceded by a sample and hold circuit. As discussed above, the hold command is generated from each second fringe of the neon signal so that data words correspond to interferogram sample points 1.1705 μ apart. The ramp converter was chosen for simplicity while maintaining sufficient accuracy. Two synchronization bits, the gain bit and a parity bit, are combined with the eight data bits to form a 12-bit word. Words are serially shifted out to one of two spacecraft tape recorders.

23-4.5 Logic and Control System

The operation of the instrument is sequenced by the logic and control system, which receives various clock signals and ground commands from the spacecraft. The logic and control circuit divides the system cycle into sixteen frames, fourteen taken while viewing the earth, and two while facing the calibration sources. Each frame is 16 sec long and has a total of 5000 12-bit words; 3408 for each interferogram, two sets of 64 words of housekeeping data bracketing each interferogram, and zeros to complete the frame.

23-4.6 Calibration and Image Motion Compensation

An image motion compensation and a calibration function (IMCC) reduce target smear and provide calibration inputs. Both functions are accomplished by rotation of the same mirror inclined at 45° to the axis of rotation as indicated in Figure 23-1. Target smear is minimized by slowly rotating the IMCC mirror while in the earth-view position through 3.5° in 10 sec, the duration of one interferogram. Each interferogram represents a signal which originates within a well-defined area of about 145-km diameter.

The automatic calibration sequence begins after the fourteenth earthviewed frame, when the IMCC mirror moves to first view deep space, and then the on-board blackbody, completing the total cycle. Image motion compensation is inhibited during the calibration process and the calibration sequence is inhibited over the north polar region, where the sun could be accidentally in the field of view during deep space calibration.

23-4.7 Power System

Two separate power supplies were required to allow heater operation with or without instrument operation. Each supply may be switched by ground command. Both supplies contain direct current to direct current converters and regulators.

23-4.8 Thermal Control

The requirement to operate the IRIS optics in orbit at 250° K and also on the ground at room temperature demanded careful consideration of problems associated with differential expansion and heat loads. The design uses three resilient mounting pads, one under each housing of the optical module. These vibration isolators between the optics housing and the support structure accomplish three objectives: they provide a flexible mechanical interface which reduces possible strains induced into the optics module; they attenuate vibration from the launch vehicle; and they increase the thermal impedance. This was not enough, however. A fiberglass hollow cylinder had to be added between the optics module and the vibration isolators to provide additional thermal isolation.

Temperature control of the optics was based on radiating heat to space, which was intended to subcool the instrument, and then replacing a portion of the heat lost with power from an electrical heater in a thermostat circuit. Radiating surfaces were selectively coated with a white paint of high emissivity in the infrared. Heaters, bonded to these surfaces, have different power consumptions to account for variations in the view factors around the curved surfaces of the optics housing. The entire instrument, except for the view ports and the cooling surface, was covered with a blanket of super insulation consisting of 32 layers of aluminized mylar. In orbit, the instrument missed its 250° K design temperature by about 5° and equalized at 255° K. Fortunately, the onboard calibration technique compensated for the effect. Furthermore, the thermal time

constant of the optics module is about 10 to 15 hours, so the temperature fluctuation was small between the sunlight and shaded part of the orbit.

23–5 DATA REDUCTION

During the interrogation of Nimbus III by the STADAN data acquisition facilities in Alaska and Rosman, N. C., interferograms were transferred from the satellite recorder to a tape recorder on the ground. The Nimbus ground station at Goddard Space Flight Center rearranged the IRIS data into computer-compatible format. The new tapes are processed on a high-speed digital computer, which first checks the data stream for completeness and processes the housekeeping information. Then all interferograms are Fourier transformed using the Cooley-Tukey method.[20,21] A smoothing or apodization function is applied to all interferograms to reduce side lobes of the instrument function. The effect of the particular apodization function used

$$(0.54+0.46\cos 2\pi\,\tau/T)$$

is discussed by Blackman and Tukey.[22] In this function τ is the distance from the center and T is the total length of the interferogram.

The Fourier transformation yields the amplitudes of the sine and cosine terms, which can be converted into magnitude and phase for each spectral element computed. The location of the pronounced central peak in the calibration interferogram is taken as the phase reference point. Atmospheric interferograms often lack a central peak, as shown in Figure 23-2, and the computer uses first a point near the expected center and derives the power spectrum and a phase curve. If the slope of the phase curve exceeds a certain amount, the phase reference point is shifted proportionally to the slope and a new amplitude and phase are computed.

The phase determination is rather important since IRIS was designed to operate at the midrange of brightness temperatures expected to be seen in orbit. The power spectrum gives only the magnitude of the difference in intensity between the source and the instrument. To resolve the ambiguity between sources colder and warmer than the instrument, the phase information must be used. The phase of a source colder than the instrument is 180° apart from the phase of a source warmer than the instrument. A typical power spectrum, a phase spectrum, and a phase-corrected power spectrum are shown in Figure 23-5. A typical uncalibrated power spectrum and the associated phase spectrum, shown in the upper and middle part of Figure 23-5, are the result of the Fourier transformation of the interferogram. The lower part of the figure shows the phase-corrected power spectrum before calibration. In the automatic

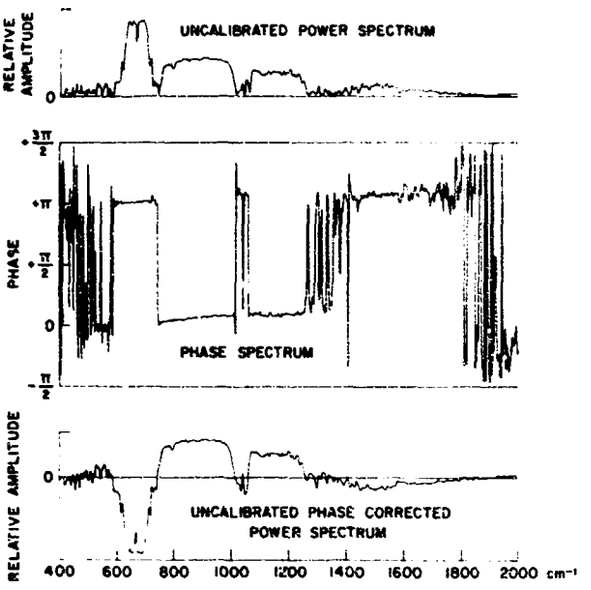


Figure 23-5. Typical Power Spectrum, Phase Spectrum, and Phase-Corrected Power Spectrum

reduction process the phase angle within the 15-μ band is taken equal to π. In that spectral interval the expected brightness temperatures are always lower than 250° K while viewing earth, and this region may therefore serve as a polarity reference. All spectra from cold blackbodies are taken to be negative, and, correspondingly, all warm calibration spectra are taken to be positive. After all power spectra have been phase corrected, that is, after each spectral component has the proper plus and minus sign attached, the actual calibration process takes place.

23–6 CALIBRATION

In-flight calibration is performed by periodically viewing an onboard blackbody of about 280° K and outer space. The calibration interferograms are transformed in the same way as the interferograms obtained while viewing earth. The amplitude C_ν in the spectrum is proportional to the difference in radiance between the target and the instrument. Thus, for each wavenumber interval,

$$C=r(B_{\text{target}}-B_{\text{instr}}). \qquad (23\text{-}2)$$

The sign of C has already been determined as discussed above. From the calibration sequence, one obtains a set of three equations: one for the target (Index t), one for the cold blackbody (Index c), and

one for the warm blackbody (Index w). The three equations may be solved to yield target brightness (B_t), responsivity (r), and instrument brightness (B_i). If one uses the interstellar background (~4° K) as the cold reference, then B_c is essentially zero and the equations simplify to

$$B_t = \frac{C_t - \overline{C}_c}{\overline{C}_w - \overline{C}_c} B_w, \tag{23-3}$$

$$r = \frac{C_w - C_c}{B_w}, \tag{23-4}$$

$$B_i = \frac{-\overline{C}_c}{\overline{C}_w - \overline{C}_c} B_w. \tag{23-5}$$

Because C_c is negative, the term $(C_w - C_c)$ is the sum of the amplitudes in the calibration spectra and not the difference between quantities of the same order of magnitude.

Equation (23-3) for B_t is used to reduce the spectra from earth. The calibration spectra $\overline{C}_w$ and $\overline{C}_c$ of Eqs. (23-3) and (23-5) are the average of all calibration spectra within one orbit. Figure 23-6 shows the orbital averages of the calibration spectra on a relative scale, the responsivity calculated from the calibration spectra, and the noise-equivalent radiance calculated from the standard deviation of the responsivity. The factor $\sqrt{2}$ of Eq. (23-7) was omitted.

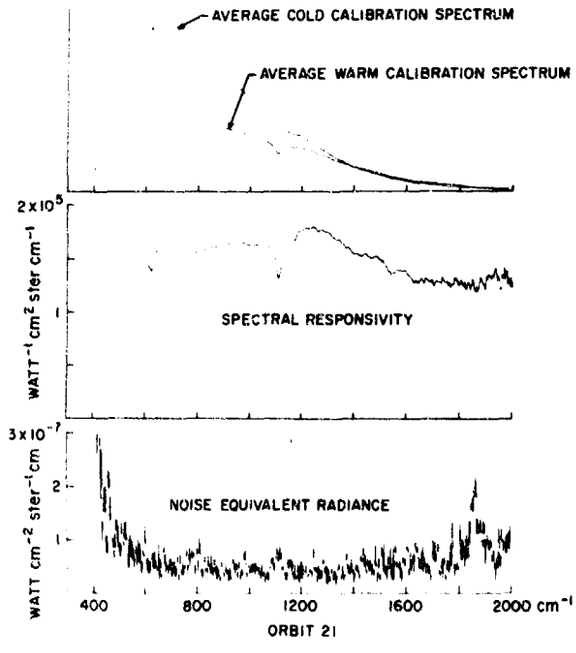


Figure 23-6. Orbital Averages of the Calibration Spectra, the Responsivity, and the Noise-Equivalent Radiance

Samples of calibration spectra and the average responsivity $\bar{r}$ are shown in the upper and central parts of Figure 23-6. A comparison of the average orbital responsivity for each spectral interval from orbit to orbit and from day to day yields the long-term drift.

The short-term repeatability of the instrument is determined by calculating the standard deviation s of the individual responsivities r_i, computed from pairs of individual hot and cold calibration spectra,

$$s = \left(\frac{\sum_{i=1}^{k} (r_i - \bar{r})^2}{k-1} \right)^{\frac{1}{2}}, \tag{23-6}$$

where k is the number of calibration pairs per orbit. The noise equivalent radiance (NER) may be calculated from

$$\mathrm{NER} = \sqrt{2}\, s B_w / \bar{r}. \tag{23-7}$$

The square root of two factor accounts for the use of two spectra in the computation of the individual responsivities. The noise equivalent radiance (NER) of the instrument in orbit is shown in the lowest part of Figure 23-6. Between turn-on of the instrument and approximately orbit 50, the responsivity decreased, and, as a consequence, the NER increased. The effect amounts to about 30 percent at longer wavelengths and more towards shorter wavelengths. Without the on-board calibration capability, the scientific objectives could not have been met. With the on-board calibration procedure the effect is compensated; however, the 1400- to 2000-cm^{-1} (5- to 7-μ) range is unreliable after orbit 50. Fortunately, this part of the spectrum is not used in the derivations of the temperature, humidity, and ozone profiles, which are necessary steps to reach the main objectives of this experiment.

23-7 INSTRUMENT PERFORMANCE

How well has the design proven itself? How close did the performance of the instrument come to the theoretically possible limit? The signal-to-noise ratio of the interferometer or of any radiometric device can be expressed by (for example, see Hanel[23])

$$S/N = \eta_1 \eta_2 I_\nu \, \Delta\nu D^* \sqrt{A \Omega \Omega_d \tau}. \tag{23-8}$$

The noise equivalent radiance is the radiance I_ν for which the signal-to-noise ratio becomes unity.

The following numerical values have been taken:

system efficiency for single detector interferometers,	$\eta_1 = 0.35$;
optical efficiency,	$\eta_2 = 0.25$;
width of resolved spectral element (cm^{-1}),	$\Delta\nu = 5$;
figure of merit for thermistor bolometer in bridge circuit [$cm(Hz)^{\frac{1}{2}} W^{-1}$],	$D^* = 10^8$;
area of non-moving interferometer mirror (cm^2),	$A = 13$;
solid angle of instrument (sr),	$\Omega = 1.6 \times 10^{-2}$;
solid angle of detector illumination (sr),	$\Omega = 1$;
Observation time per interferogram (sec),	$\tau = 10$.

The optical efficiency includes the reflection losses on three mirrors, 0.95 each; the transmission losses on the entrance window, 0.7; obscuration, 0.7, by the reference interferometer and spiders; and finally, the efficiency of the beamsplitter, 0.6. The calculated noise equivalent radiance becomes 1.6×10^{-8}. The observed NER is about 3 to 5 times higher. Most of the difference is accounted for by the noise introduced by the bolometer power supply and by imperfect alignment. The preamplifier and the quantization process also contribute small amounts of noise. The performance of the instrument came fairly close to the theoretically possible limit; however, there is stil' room for improvement.

23-8 RESULTS

A typical atmospheric emission spectrum taken over a clear tropical ocean is shown in Figure 23-7.

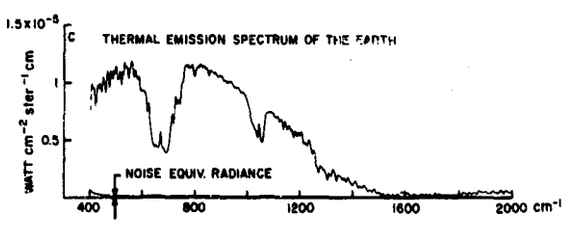


Figure 23-7. Typical Emission Spectrum of a Tropical Region

The strong CO_2 bands between 500 and 780 cm^{-1} were used to derive the atmospheric temperature profile. Ozone absorption appears between 1000 and 1070 cm^{-1}. Most other features are due to H_2O except for a CH_4 band near 1306 cm^{-1} and weak bands of CO_2 and N_2O. The spectrum was calibrated by the procedure outlined above. The dominant spectral feature between approximately 600 and 750 cm^{-1} is due to the CO_2 molecule. The higher intensities in the more transparent wings of the band reflect warmer temperatures of the lower layers in the atmosphere, and the lower intensities toward the band center correspond to the lower temperatures near the tropopause. The most opaque part of the band, the Q-branch near 667 cm^{-1}, indicates an opposing trend caused by higher stratospheric temperatures. Radiances in the CO_2 band are used to derive atmospheric temperature profiles.

If the temperature profile is known, it is possible to deduce the distribution of other atmospheric constituents such as O_3 and H_2O. The ozone band centered at approximately 1040 cm^{-1} and water vapor absorption regions between 1200 and 1400 cm^{-1} and between 400 and 550 cm^{-1} are well suited for that purpose. Between 800 and 1000 cm^{-1} and between 1070 and 1250 cm^{-1} the radiation originates predominantly from the surface, with some absorption by water vapor and carbon dioxide. For example, this spectral interval can be used to determine the surface brightness and, in many cases, the surface temperature. The same figure gives the noise equivalent radiance computed from the calibration interferograms of that orbit. Figure 23-8 illustrates the

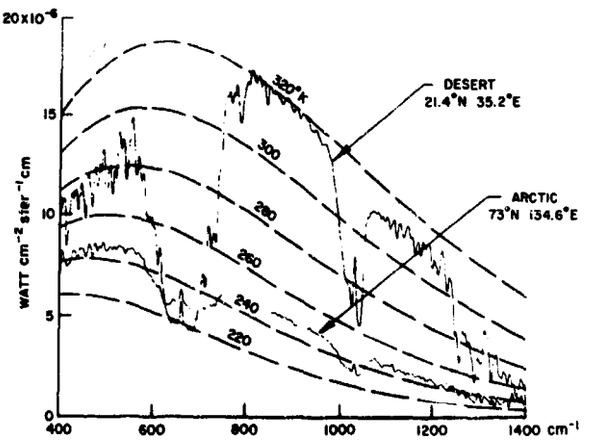


Figure 23-8. Spectra Recorded by IRIS Over Desert and Arctic Regions

extreme difference between desert and arctic spectra. The desert soil at noon is warmer than the air and shows a reststrahlen effect, a lower emissivity in the 1100-cm^{-1} region compared with the 900-cm^{-1} region, typical of silicate oxides. Brightness temperatures of blackbody radiation are also shown. A discussion of the spectra and an interpretation of the data will be published elsewhere.[24]

23-9 SUMMARY

At the beginning of this project, it seemed almost beyond the state-of-the art to record the infrared emission spectrum over a wide spectral range (400 to 2000 cm^{-1}) with a resolution equivalent to 5 cm^{-1} and with a precision of 1 percent from a satellite. The requirement of a long orbital lifetime limited the design to a thermistor bolometer, which is a relatively poor detector compared with cryogenically cooled devices. Only the large area times solid angle and the multiplex advantages of the Michelson interferometer promised a possible solution, but the performance of the interferometer had to be close to calculated expectations. The instrument discussed in detail in this paper achieved the required performance and came close to the calculated limits of resolution and precision. The main accomplishments of the design were the maintenance of adequate optical alignment over a 50° C temperature range; the operation of the moving mirror of the interferometer by a phase-locked loop to obtain the required radiometric precision of about 1 percent; and finally, the achievement of the required accuracy by an on-board calibration technique.

In orbit the instrument performed well, although not flawlessly. The spectra obtained were more than adequate for the achievement of the scientific objectives of the experiment. The desired equilibrium temperature was not reached, however, and some misalignment caused degradations primarily at the short-wavelength end of the spectrum. Nevertheless, it was demonstrated that such a delicate instrument as a Michelson interferometer can be flown on a spacecraft and produce spectra of good quality.

Acknowledgments

The authors are grateful to B. Conrath, D. Rogers, and L. Watson for helpful comments concerning the manuscript.

References

1. Fellgett, P. (1951) Thesis, Univ. of Cambridge.
2. Jacquinot, P., and Dufour, C. (1948) *J. Rech. du Centre Nat. Rech. Sci. Lab. Bellevue (Paris)* **6**: 91.
3. Jacquinot, P. (1960) Reports on progress in physics, *Phys. Soc., London* **23**: 267.
4. Connes, J. (1961) *Rev. Opt.* **40**: 45, 116, 171, 231.
5. Mertz, L. (1965) *Transformation in Optics*, John Wiley, New York.
6. Loewenstein, E. V. (1966) *Appl. Opt.* **5**: 5, 845.
7. Vanasse, G., and Sakai, H. (1967) *Progress in Optics* (6), North-Holland Publishing Co., Amsterdam.
8. King, J. I. F. (1958), in *Scientific Users of Earth Satellites*, J. A. Van Allen, Ed., Univ of Michigan Press.
9. Kaplan, L. (1959) *J. Opt. Soc. Am.* **49**: 10, 1004.
10. Wark, D., and Fleming (1966) *Monthly Weather Rev.* **94**: 351.
11. King, J. I. F. (1964) *J. Atm. Sci.* **21**: 324.
12. Chahine, M. T. (1968) *J. Opt. Soc. Am.* **58**: 1634.
13. Conrath, B. J. (1969) *J. Geophys. Res.* **74**: 3347.
14. Mertz, L. (1963) *Proc. 12th Intern. Astrophys. Symp.*, Liege, p. 120.
15. Block, L. C., and Zachor, A. S. (1964) *Appl. Opt.* **3**: 209.
16. Lovett, J., Marcotte, L., and Nadile, R. (1967) *Gemini V, D4/D7 Spectral Measurements of Space Objects and Earth-Cloud Backgrounds*, AFCRL-67-0563.
17. Chaney, L. W., Drayson, S. R., and Young, C. (1967) *Appl. Opt.* **6**: 347.
18. Hanel, R. A., and Chaney, L. W. (1966) *Proc. XVII Intern. Astron. Congr.*, Madrid, Vol. 2.
19. Hanel, R. A., and Conrath, B. J. (1969) *Science* **165**: 1258.
20. Cooley, J. W., and Tukey, J. W. (1965) *Mathematics of Computation* **19**: 296.
21. Forman, M. L. (1966) *J. Opt. Soc. Am.* **56**: 978.
22. Blackman, R. B., and Tukey, J. W. (1958) *The Measurement of Power Spectra*, Dover Publications, Inc., New York.
23. Hanel, R. A. (1969) Recent advances in satellite radiation measurements, Advances in Geophysics, Vol 14.
24. Conrath, B., Hanel, R. A., and Prabhakara, C. (1970). To be published.

Discussion

Q. (G. M. Levy, Barringer Research Ltd.): Have you profiled water vapor and ozone?

A. (R. A. Hanel): Yes.

Q. (G. M. Levy): Have you profiled any other gases?

A. (R. A. Hanel): No.

Q. (J. D. Strong, University of Massachusetts): What was the origin of the irregularities in the cold and warm calibration spectra?

A. (R. A. Hanel): Absorption in the beamsplitter coatings.

Q. (Robert J. Bell, University of Missouri): What is the lifetime of the instrument in flight?

A. (R. A. Hanel): About $3\frac{1}{2}$ months; the thermister burned out.

Contents

24. Michelson Interferometer-Spectrometer for On-Board Space Vehicle Measurements

G. F. Hohnstreiter and Thomas P. Sheahen
Bell Telephone Laboratories, Inc.
Whippany, New Jersey

W. Howell
Block Engineering, Inc.
Cambridge, Massachusetts

Abstract

A rugged lightweight Michelson interferometer-spectrometer was developed and flown successfully aboard a space vehicle. Two interferograms were produced. A Si detector was used between 9000 and 11,400 cm^{-1}, and a thermoelectrically cooled PbSe detector was used between 1800 and 9000 cm^{-1}. Both had a dynamic range of 10^6, noise equivalent signal input of 5×10^{-6} W/cm^2 sr μ, 10 interferograms/second, and a spectral resolution of 20 cm^{-1}.

24-1 INTRODUCTION

A series of Michelson interferometer-spectrometers have been designed, developed, and flown successfully on several space vehicles. Ten flight instruments were developed and built by Block Engineering, Inc., under the direction of Bell Telephone Laboratories, Inc., for infrared flight measurements. The interferometer-spectrometer was well suited to space vehicle measurements since no instrument of comparable size and weight could equal both its sensitivity and its resolution over the required spectral range. The vehicle operating environment required a rugged instrument, insensitive to vibration and acceleration. The available power was limited, and the interferometer had to share a specially designed analog data link with other instruments.

The basic characteristics of the instrument were a wavelength range of 0.86 to 5.3 μ, a resolution of 20 cm^{-1}, a scan rate of 10 interferograms/sec, and typical sensitivities of 5×10^{-6} W/cm^2 sr μ. In this paper the flight instrument and associated components are described. The flight telemetry and the data reduction methods used in transforming the interferograms to spectra are discussed in Chapter 25.

The basic principles of Fourier transform spectroscopy will be omitted here because they have been well covered elsewhere.[1-7] In Section 24-2 a discussion of the criteria governing the design of the interferometer is given. Section 24-3 contains a description of the instrument.

It was found necessary to chirp the instrument in order to provide acceptable signal-to-noise ratios in

the telemetered interferogram. The chirping methods are discussed in Section 24-4. The instruments were subjected to a thorough calibration prior to launch. The calibrations were performed independently at two separate laboratories with generally excellent agreement. Some results of the calibrations are discussed in Section 24-5. Section 24-6 contains an abstract of the data-link, and Section 24-7 the conclusions of this paper.

24-2 DESIGN CONSIDERATIONS

The instrument design was determined by both scientific requirements on data quality and hardware constraints imposed by the vehicle and operating environment.

The design requirements for the flight interferometer were: high sensitivity over a spectral range of 0.86 to 5.3 μ, high scan or repetition rate, moderate spectral resolution, at least 10^6 dynamic range in radiance, and a wide field-of-view. In addition, the instrument had to remain insensitive during operation to, collectively: vibration of amplitude 11.6 g rms, accelerations up to 40 g, and a temperature range of -25 to $+55°$ C. It had to accomplish these factors not only in a vacuum environment but at all pressures (continuously varied) up to atmospheric pressure. Moreover, the interferometer had to survive a series of 100-g shock tests. The physical size constraints were $(1/6 \times 1/4 \times 1/2)$ m^3, the weight limit was 17 kg, and the maximum power available for operation was 100 W. One dominant constraint was the availability of only 40 dB of telemetry dynamic range. Since an overall dynamic range of 10^6 (120 dB) in the source intensity was expected, automatic gain switching of the interferometer was required.

The spectral resolution was limited to 20 cm^{-1} by the required scan rate and the sensitivity requirements. The available signal-to-noise ratio did not justify additional resolution. Wavenumber accuracy of 10 cm^{-1} represents the typical noise-limited case. The basic wavenumber scale is linear, and is directly relatable to an argon line used in the reference interferometer.

The required spectral range was covered in two channels, which allowed selection of the optimum detector type for each region. In addition, the sharing of the interferogram dynamic range burden by two channels resulted in improved data quality.

Optimizing the design goals and constraints led to the operating characteristics listed in Table 24-1. Figure 24-1 is a photograph of an actual flight unit.

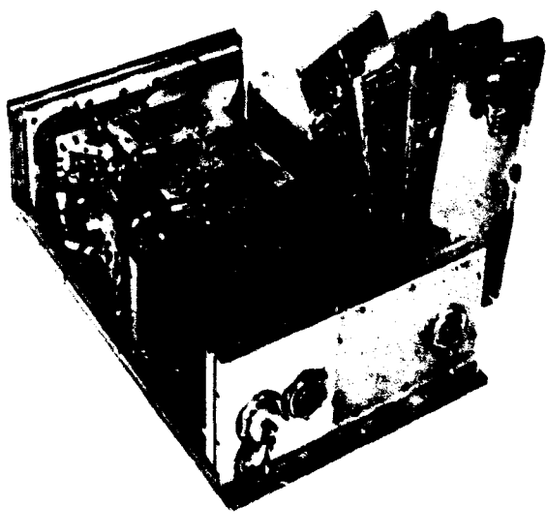


Figure 24-1. Flight Model Interferometer

24-3 INSTRUMENT DESCRIPTION

The Michelson interferometer provides spectral coverage from 0.86 to 5.3 μ by incorporating two infrared detectors behind a single interferometer.

A dichroic thin-film at the exit of the interferometer directs the energy to the appropriate detector. Each detector is coupled to an independent signal-processing channel.

Table 24-1. Michelson Interferometer-Spectrometer Operational Specifications

A. *Scientific Properties*	*Channel 1*	*Channel 2*
Spectral Regions, μ	1.1–5.3	0.86–1.1
Sensitivity (NESR), W/cm^2 sr μ	5×10^{-6} @ 4 μ	5×10^{-6} @ 1 μ
Detectors	PbSe (−50° C)	Si
In-flight Wavelength Calibration Line:		
Wavelength μ	None	He 1.083
Reference Line	8521 Å Argon	
Optical Retardation, μ	870	
Dynamic Range	10^6	
Spectral Resolution	20 cm^{-1}	
Wavenumber Accuracy	10 cm^{-1}	
B. *Mechanical Properties*		
Instrument Type	Fringe referenced double cube	
Drive System	Electromagnetic, servo controlled	
Scan Rate	10 scans/sec	
Duty Cycle	80%	
Mirror Diameter	11 mm	
Field of view (conical)	35° (PbSe instrument)	

C. *Electrical Properties*	
Number of Gain States per Channel	8
Gain State Indication	Equal Voltage Steps, 0–5 Vdc
Telemetry Dynamic Range	40 dB
Signal Bandwidth (Channel 1)	1220 Hz to 9000 Hz
Signal Bandwidth (Channel 2)	6100 Hz to 14,350 Hz
Signal Bandwidth (Reference)	22,000 Hz to 28,700 Hz
Signal Output	0 to +5 V
Maximum Output Limits	−0.6 V, +6.0 V
Output Impedance	300 Ω, 1500 pF
Supply Voltage	27.5±1 Vdc
Average Power Consumption	60 W
D. *Physical Characteristics*	
Environmental Specification Temperature Range	−25 to +55° C
Expected Operating Temperature Ranges:	
Instrument Ambient	+5 to +40° C
PbSe Detector	−50° C
Vibration Level	11.6 g rms, 0 to 2000 Hz
External Dimensions	meters: 0.127 H×0.264 W (window end)×0.422 L inches: 5.0 H×10.4 W (window end)×16.62 L
Maximum Weight	17 kg=37 lb

The movable mirror was continuously scanned to a maximum path difference of 870 μ at a repetition rate of 10 scans per sec. Continuous scanning resulted in a direct coding of wavenumber in terms of electrical frequency, which facilitated analog telemetry. The minimum frequency of about 1 kHz was above major source fluctuation frequencies, thus eliminating excess noise from the spectra. This is the technique first used by Mertz[6] to eliminate atmospheric scintillation effects in astronomical spectra. The frequency range listed in Table 24-1 includes excess bandwidth to accommodate random frequency perturbations caused by vibration.

24–3.1 Optical System

The optical system is shown in Figure 24-2. It consists essentially of two interferometer cubes whose moving mirrors are interconnected by a rigid shaft.

To achieve 20-cm^{-1} resolution at the short wavelengths, energy from the 35° field is collimated to 1.5° by lenses L1 and L2 before it enters the signal interferometer. The modulated output beam is reconverged by lens L3 and directed by a dichroic beamsplitter to the silicon and lead selenide detectors through fast lenses L4 and L5, which focus the beams on the smallest possible detectors.

The silicon diode on the signal interferometer provides the instrument response between 0.86 and 1.1 μ and is operated at ambient temperature. From 1.1 to 5.3 μ the instrument response is provided by the PbSe detector cooled to −50° C. For this purpose a four-stage thermoelectric cooler is used to provide a 95° C temperature differential between the detector and the cooler base. A thermistor mounted on the

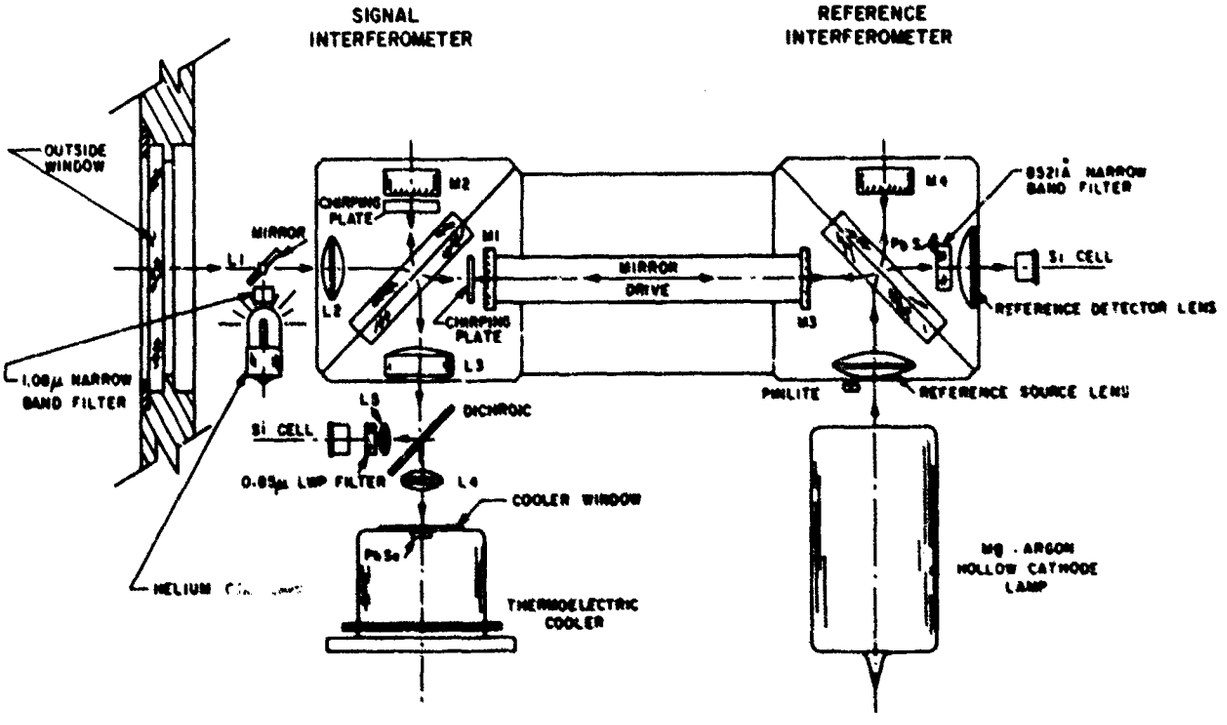


Figure 24-2. Schematic Diagram of the Optical System

top stage of the cooler provides feedback to a proportional temperature controller, which controls the detector temperature to ±1° C. It was found that some lead selenide detectors exhibited significant nonlinearity at high signal levels, but individual selection of detectors reduced this effect. In order to prevent spectral aliasing, an 8520 Å long wave-pass filter was placed over the silicon detector.

The interferograms were chirped by 0.375 cm of Irtran 5 compensated by 0.602 cm of CaF_2. A discussion of chirping and compensation requirements is given by Mertz.[6]

Inflight calibration of the instrument is provided only on the short-wavelength channel. A helium discharge lamp is mounted near the entrance aperture of the instrument. The calibration lamp is always on, so that it produces a continuous calibration. Energy from the 1.083-μ He line is passed through a narrow-bandpass interference filter. The radiation is introduced into the signal interferometer in sufficient intensity to produce a calibration line with a 100:1 signal-to-noise ratio when the automatic attenuator is in the highest gain state. As the attenuator reduces the channel gain, the signal-to-noise ratio of the calibration line is proportionally reduced until the line is eventually obliterated by the received spectra.

The position of the moving mirror is controlled by a mechanically slaved second interferometer cube. The reference interferometer continuously monitors the position of the moving mirror.

The purpose of the servomechanism is to maintain the scan linearity within tolerance while the instrument is subjected to the vibrations and accelerations of the space vehicle. The servomechanism also corrects the thermal dependency of the transducer. The error signal for the servo is produced by the reference interferometer. The moveable mirrors of both signal and reference interferometers are driven by the common transducer. With its silicon detector, the reference interferometer views a Mg-Ar hollow cathode lamp filtered to pass only the 8521-Å line. The reference interferogram is virtually sinusoidal. Its frequency is proportional to mirror velocity. Comparison of this frequency to a set frequency yields the error signal for the servo. In order to indicate the beginning of each scan, a PbS detector in the reference cube observes a white-light fringe from an incandescent lamp. The use of a separate interferometer for this purpose facilitates beamsplitter optimization, and eliminates the possibility of scattered light from the reference sources entering the signal beam. The separate interferometer also prevents crowding of components around the main interferometer.

24-3.2 Drive System

The moving mirror is driven by an electromagnetic transducer working against springs. In the absence of external acceleration and vibration, a constant velocity is obtained by varying the drive current linearly with time. Velocity variations produce frequency modulation of both the data and reference interferograms. Whereas small amounts of FM are correctable by fringe-reference sampling, large amounts are not, because the reference interferogram might lie outside the data-link frequency bandpass.

The susceptibility of the transducer to vibration is essentially governed by the available drive power. For the vibration levels anticipated, a transducer designed for full available power would suffer large velocity errors, large enough in fact to reverse the motion in midsweep. Figure 24-3 shows velocity

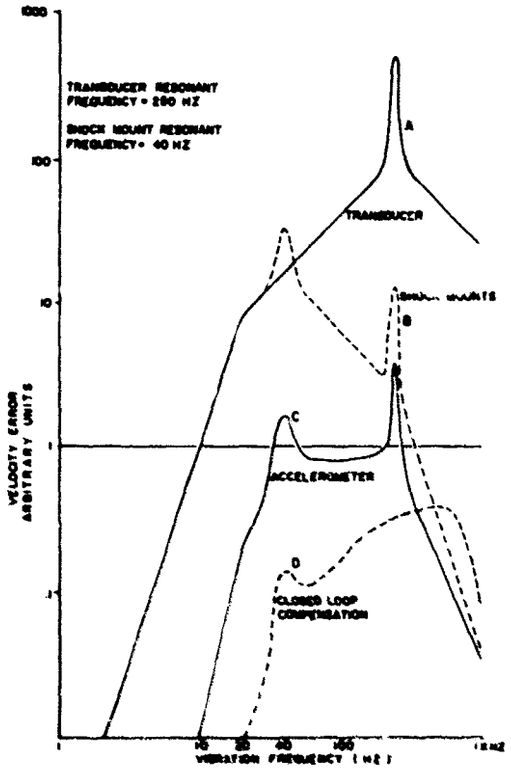


Figure 24-3. Velocity Error as a Function of Vibration Frequency for the Design Vibration Level

error as a function of vibration frequency for the design vibration level. Curve A represents the performance of the transducer alone. The successive curves show the cumulative effect of the velocity stabilization measures applied. Curve B shows the improvement due to shock mounting the interferometer. Large reductions in velocity error occur for vibration frequencies above the resonant frequency of the shock mounts. The minimum shock-mount resonant frequency was governed by the requirement that the mounts should not bottom under the specified dc acceleration. Curve C shows the effect of modifying the driving waveform according to the output of an accelerometer, mounted parallel to the transducer axis, which senses the vibration transmitted through the shock mounts. Curve D is the result of a closed-loop servo controlling the mirror motion. The error

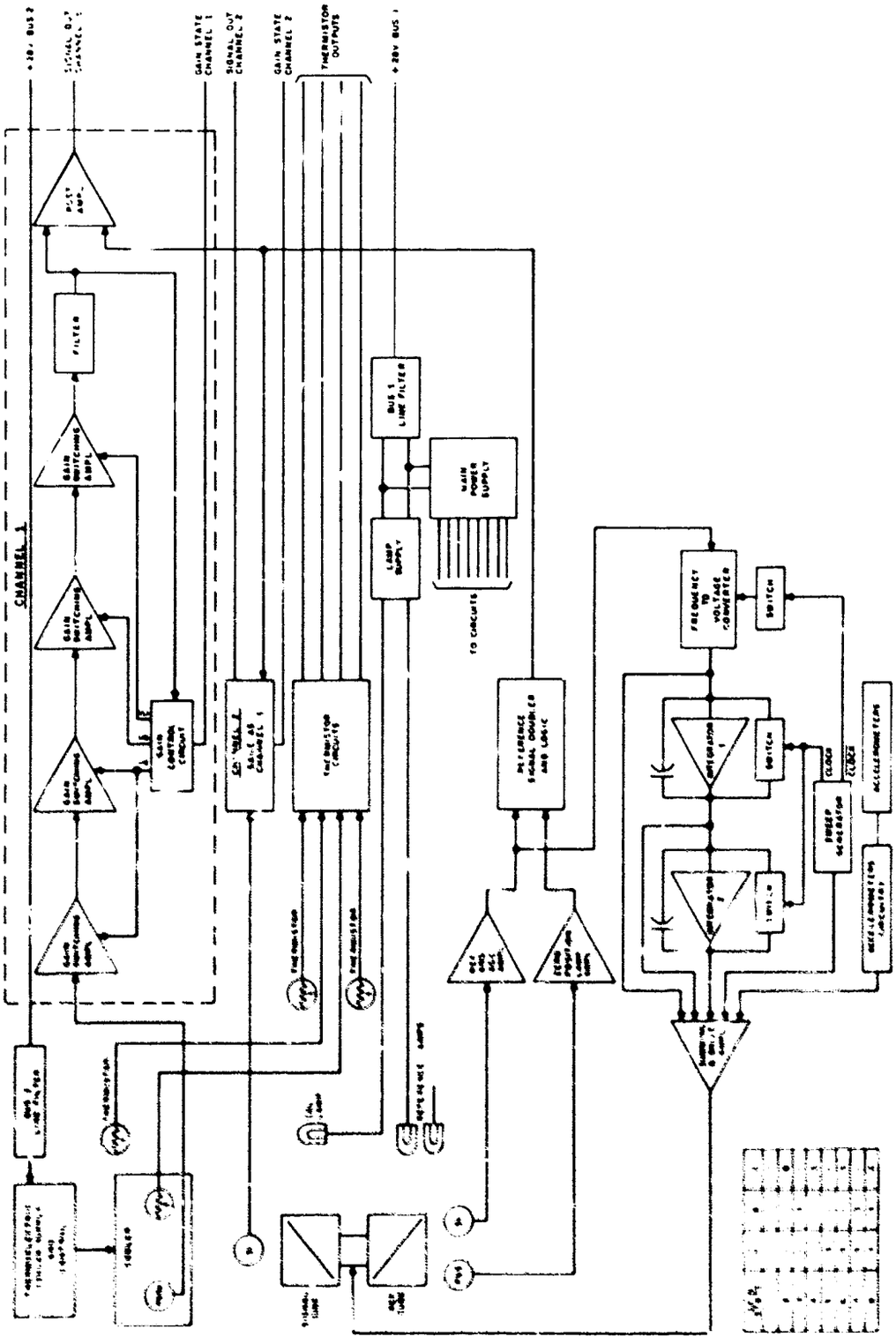


Figure 24-4. A Block Diagram of Instrument Electronics

signal is derived by subtracting the frequency of the monochromatic reference interferogram from a fixed electrical reference. The closed-loop compensation includes velocity feedback for increased damping, position feedback for increased resonant frequency, and integral of position feedback for added attenuation of vibration effects. The resulting velocity is sufficiently uniform to allow fringe-referenced digitization with the received data; however, it is not sufficiently linear to permit digital sampling of the data interferogram on a constant time base. Therefore, in addition to providing an error signal to the servomechanism, the reference interferogram is also used for digitizing the data interferogram.

24-3.3 Electronics

Functionally, the electronics in the interferometer can be divided into the following areas: transducer driver circuitry, reference signal circuits, and signal processing circuitry. A block diagram is shown in Figure 24-4.

The transducer consists of two coils, rigidly coupled to a drive shaft that moves in a magnetic field derived from a permanent magnet. The drive shaft and coil assembly are suspended in the magnetic field by means of two flat springs. In principle the coils are driven by a sawtooth waveform centered around 0 V, thus sweeping the mirror at a constant velocity. The sweep generator consists of an operational amplifier connected in an integrator configuration. A small current fed to the input of the amplifier causes its output voltage to rise at a uniform rate. A comparator senses the output voltage level. As soon as the ramp voltage reaches a predetermined value a monostable circuit is triggered, which resets the integrator to zero, terminating the scan. The sweep generator output is summed with servo error signals, amplified, and presented to the coils of the transducer.

The error signals are derived from the monochromatic interferogram from the reference interferometer. After appropriate amplification and automatic gain control, this sinusoidal signal is presented to a frequency-to-voltage converter. The output of the frequency-to-voltage converter is first smoothed and amplified, and then compared to a fixed voltage. The difference voltage, if any, is a measure of the deviation of the mirror velocity from the desired mirror speed. This signal, appropriately scaled, is one of the error signals added in the summing amplifier. Since the transducer-interferometer combination has a complex transfer characteristic, additional compensation is necessary. This takes the form of two integrators in series, the output of each are applied to the summing amplifier as error signals.

The sine-wave interferogram generated by the reference interferometer is also used for sampling. It is frequency-doubled and added to the signal interferometer outputs to assure that interferogram and reference are identically signal-processed. The reference is separated by filtering on the ground and is used to command the digitizing of the signal interferogram. This fringe-referenced digitizing occurs at each positive-going zero crossing of the reference signal, allowing the shortest wavelength component of the interferogram to be 8521 Å for no aliasing.

The two signal channels are identical, with the exception of the preamplifiers. Channel 1 uses a thermoelectrically cooled PbSe detector. Near the detector are two thermistors. One is used to control the detector temperature via a series regulator preceding the converter that powered the thermoelectric cooler. The other is used to telemeter the detector temperature. Channel 2 has an ambient temperature silicon diode detector.

24-3.4 Dynamic Range

Both detectors are biased from variable power supplies. The bias is adjusted in accordance with the amount of source radiation. This feature is required to span the large dynamic range. The signal channels consist of four amplifiers in series. Each of these amplifiers can be switched in gain. In parallel with the normal high gain feedback resistor is a switchable lower gain feedback resistor. The switching is accomplished by turning on a field effect transistor in series with the low-gain feedback resistor.

To provide the 10^6 operating dynamic range of the flight interferometer, eight levels of automatic gain changing are required (using an automatic attenuator) to keep the signal levels within the available telemetry range of 0 to 5 volts. Gain is changed between interferograms by factors of 3.8, depending on whether the previous interferogram exceeds a maximum or minimum threshold value. The maximum value of each interferogram is compared with preset switch-down and switch-up levels. If either level is reached, the automatic attenuator will switch one step in the appropriate direction before the next scan begins. Seven scans are required for the attenuator to traverse its full range because, on any one scan, only too-high, correct-gain, and too-low information are available. For increasing source intensities the detector noise is soon reduced below the telemetry noise, thereby producing a signal-to-noise ratio dependent only on telemetry noise and the interferogram voltage. An increase in source intensity would produce a corresponding increase in the final spectrum signal-to-noise ratio up to the point where an additional attenuation step is switched into the circuit.

24-4 CHIRPING

In many interferometric systems, the noise appearing in the final data depends mainly on the limitation of dynamic range of the instrumentation and data processing system used. Because each wavelength has zero phase shift at the same point, in a conventional interferometer almost all interferograms have large central fringes. Therefore, because of a lack of sufficient dynamic range in the telemetry and recording system, it is often difficult to provide acceptable signal-to-noise ratios that cover the range required for subsequent Fourier transform processing of the interferogram.

It is possible, however, to modify the signal inter-

ferogram such that each wavelength has its own center, thus avoiding the large central fringes that normally occur. This distribution of wavelength centers is accomplished by placing a flat optical element into each beam before the point of beam recombination, thereby shifting the position of zero-order interference. If one of the plates has dispersion in the group velocity refractive index, then each wavelength will have its center shifted by a different amount and the interferogram is said to be "chirped".[6] There is no single path difference at which all wavelengths add in-phase to produce a large-amplitude white-light fringe. The white-light fringe is essentially dispersed over a portion of the interferogram. Since only phase is involved, the interferogram power is unaltered.

Figure 24-5 shows the dispersion in white-light fringe location for two possible chirping configurations. Although Irtran 4 produces the greater average dispersion, the Irtran 5 configuration is used since it produces maximum dispersion near the peak response of the PbSe detector. As can be seen, the white-light fringe for the first 2000 wavenumbers is spread over 500 μ of path difference. The frequencies above 10,000 cm^{-1} overlap the low frequencies somewhat, but these are received by the silicon detector and telemetered separately. Figure 24-6 shows a typical laboratory interferogram and its corresponding spectrum. As expected, the low frequencies are spread over a wide range, and the high frequencies begin to converge toward a single point. On subsequent

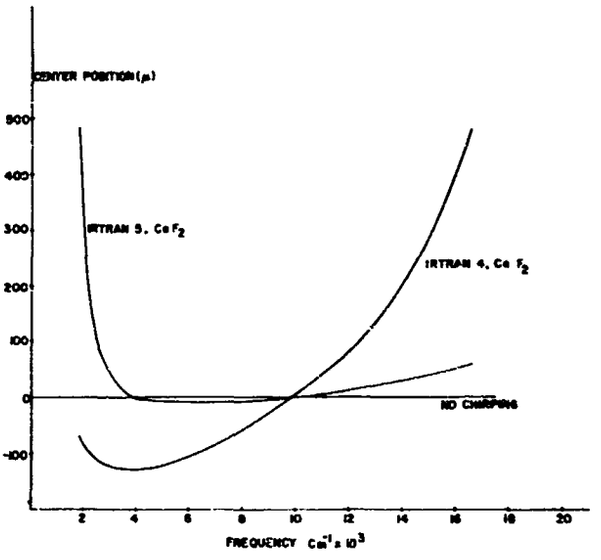


Figure 24-5. Dispersion in White-Light Fringe Location for Two Possible Chirping Configurations

reduction of the interferogram to an optical spectrum, the shifting of the center with respect to wavelength can be ignored, as long as no power is lost in clipping the wings. When power is clipped, a wavelength-dependent apodization function is required.

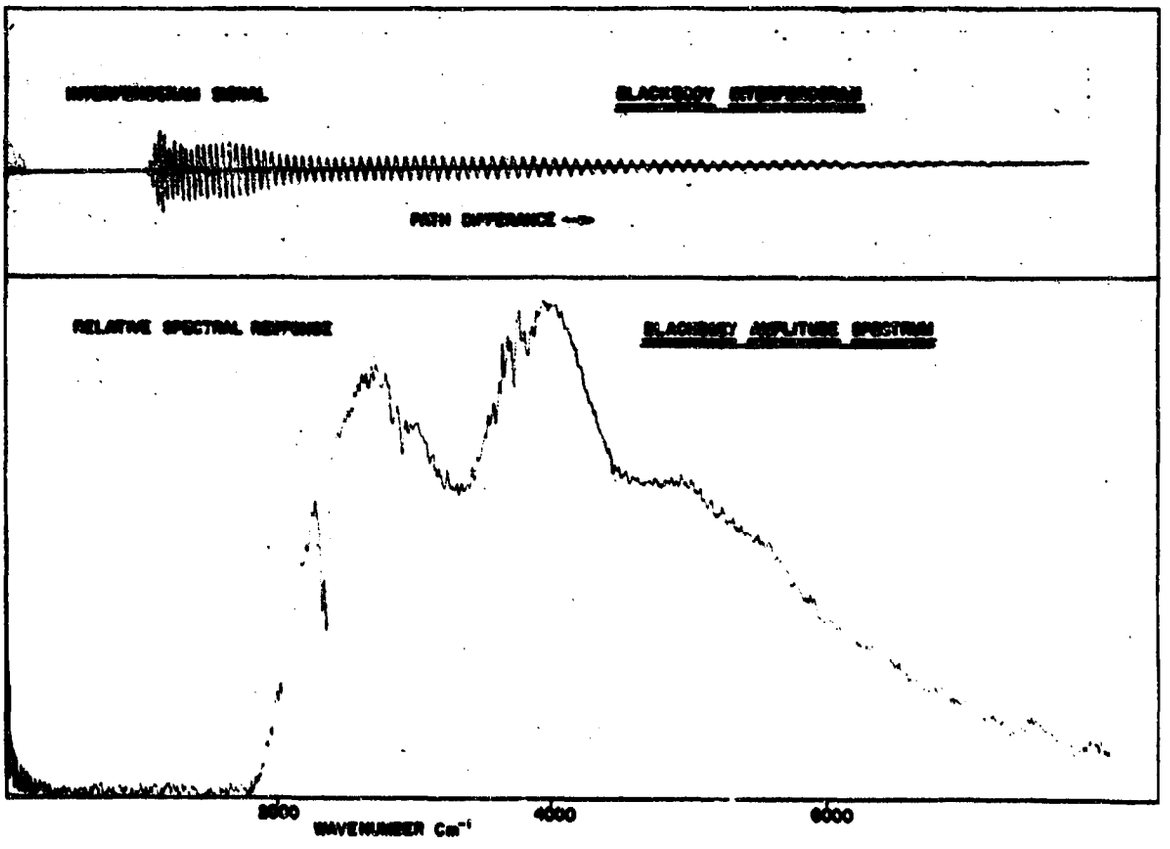


Figure 24-6. Typical Interferogram and Corresponding Spectrum

24-5 TESTS AND CALIBRATIONS

The test procedures used for the Michelson interferometer were designed to verify the ability of the various instruments to perform within design tolerances and specifications. The instrument is thoroughly tested environmentally.

The calibrations obtained quantitative information essential to subsequent data reduction and analysis. The calibration procedures utilized for the flight-model interferometer are in many cases similar to the procedures used for conventional grating spectrometers. The instruments were calibrated for the effects of field-of-view, wavelength accuracy, wavenumber resolution, instrument sensitivity, responsivity, and related effects. All calibrations were performed independently by two laboratories (Block Engineering and TRW Systems) to ensure validity of the calibration results.

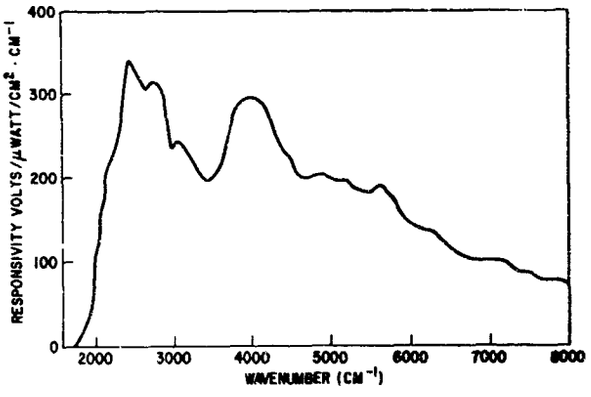


Figure 24-7. PbSe Channel Spectral Responsivity

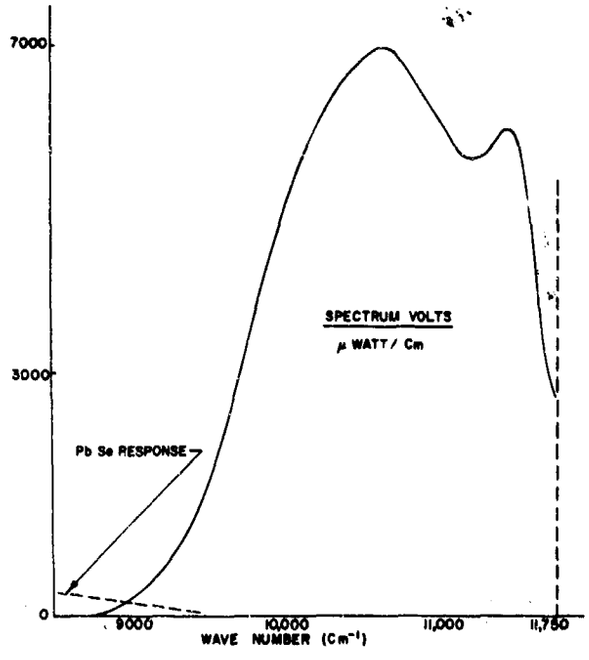


Figure 24-8. Si Channel Spectral Responsivity

24-5.1 Responsivity

Absolute spectral responsivity and sensitivity is achieved through the use of quartz-iodine irradiance standards and blackbody sources, in much the same way as for conventional spectroscopic instruments. One difference is the fact that spectral averaging of many scans is supplanted by co-adding of many interferograms. Such co-adding must be done coherently, or else a seeming reduction in responsivity will occur.

Typical values of spectral response of the instrument are shown in Figure 24-7 for the PbSe channel and in Figure 24-8 for the Si channel. Figures 24-9 and 24-10 show the relative response of the instrument in each of the eight gain-states for the PbSe channel and the Si channel.

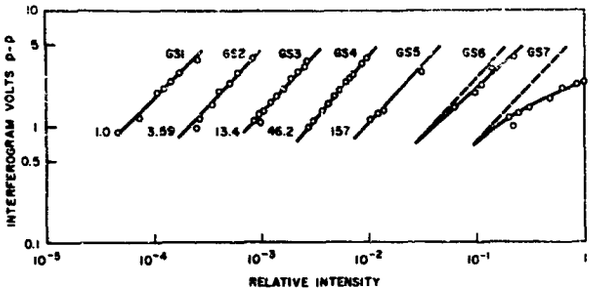


Figure 24-9. PbSe Channel Absolute Responsivity

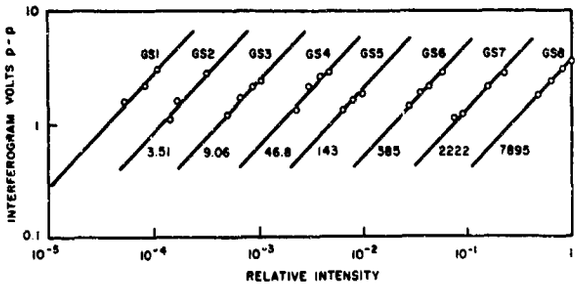


Figure 24-10. Si Channel Absolute Responsivity

24-5.2 Field-of-View Calibration

By using a spectrally narrow light source collimated by a long mirror, the entrance aperture of the interferometer is uniformly filled by a beam of less than ±1/2-degree divergence. The instrument, during calibration, is mounted on a two-axis rotary table with its entrance aperture at the center of rotation of the table. The interferometer output is recorded continuously as it is rotated in angular position in

the horizontal and vertical planes, covering ±90° in each axial direction. By "rolling" the instrument 45° about its optical axis a second set of two scans is recorded. Reduction of the resultant interferograms provides information from which an eight-point polar plot of the responsivity is plotted. Equal response contours are plotted for several wavelengths at 90, 75, 50, 25, 10, and 1 percent of peak responsivity. The scan of ±90° on each axis assures that no spurious sidelobes or irregularities, within the 10^3 dynamic range per scan, exist in the field of view. Figure 24-11 shows a typical field-of-view map for the PbSe channel.

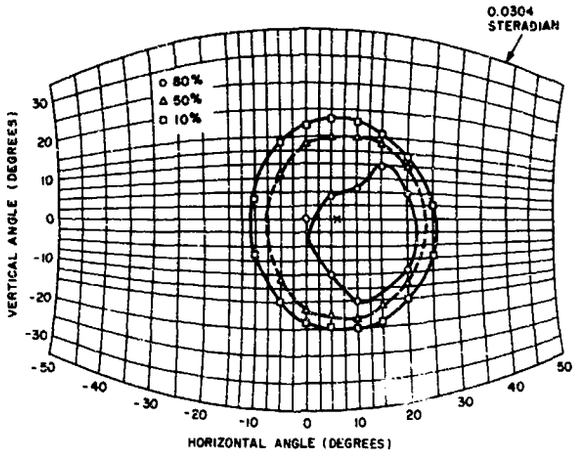


Figure 24-11. Typical Field-of-View Contours

24-5.3 Wavelength Calibration and Resolution

The wavelength calibration is performed with a source producing lines of well-known wavelengths. The instrument is calibrated against three or four spectral lines over its wavelength range. The resolution is also measured at several different positions within the field-of-view of the instrument.

24-5.4 Calibration Errors

The responsivity determined in calibration has a certain fractional uncertainty $\Delta R/R$, which is accumulated from many sources. The calibration for the lead selenide channel was obtained using a blackbody source, so the responsivity is

$$R(\omega)=\frac{Mp(\omega)\tau\left[\exp\left(\frac{1.44\omega}{T}\right)-1\right]}{K\omega^3}\cdot\frac{d^2}{A_s}, \qquad (24\text{-}1)$$

where R = responsivity, in V/(W/cm^2-cm^{-1}), M = maximum spectral amplitude in V, p = fractional value of maximum at ω, ω = frequency, τ = window transmission, K = constant factors, T = temperature, d = distance of blackbody from instrument aperture, and A_s = area subtended by blackbody source. The mean-squared error is therefore,

$$\left(\frac{\Delta R}{R}\right)^2=\left(\frac{\Delta M}{M}\right)^2+\left(\frac{\Delta p}{p}\right)^2+\left(\frac{\Delta\tau}{\tau}\right)^2+\left(\frac{2\,\Delta d}{d}\right)^2+\left(\frac{\Delta A_s}{A_s}\right)^2+\left(\frac{3\,\Delta\omega}{\omega}\right)^2+\left(\frac{1.44\exp\left[\frac{1.44\omega}{T}\right]}{T^2\left\{\exp\left[\frac{1.44\omega}{T}\right]-1\right\}}\right)^2\times([T\,\Delta\omega]^2+[\omega\,\Delta T]^2). \qquad (24\text{-}2)$$

Despite the complexity of Eq. (24-2), it reduces rapidly to a few terms, since errors in p, A_s, and ω are negligible, and errors in τ and T can be ignored compared with errors in M and d. The error in d is 4 percent, because the blackbody source is placed very close to the window; the position of the interferometer relative to the window is very difficult to measure accurately in the large vehicle, and so the distance d contains an uncertainty of 4 mm; and the smallness of d inflates the percent error in measuring d. Thus, $2\,\Delta d/d = 8$ percent.

The error in maximum spectral amplitude M amounts to about 11 percent, owing chiefly to the digitizing and calculating process for calibration data. It must be remembered that the signal originating at the interferometer passes through the data reduction system in a complete calibration which simulates the flight configuration. The instrument output is the superposition of the data interferogram with the reference square wave, which is played back through assorted filters and digitized. Further explanation of this error source is given in Chapter 25.

For the net error in responsivity,

$$\left(\frac{\Delta R}{R}\right)^2\approx\left(\frac{\Delta M}{M}\right)^2+\left(\frac{2\,\Delta d}{d}\right)^2+\text{small terms}, \qquad (24\text{-}3)$$

which gives $\Delta R/R \approx 14$ to 15 percent as the statistical accuracy of calibration.

The calibration of the silicon channel is slightly more accurate. For an Eppley quartz-iodine standard lamp of spectral irradiance $\epsilon(\omega)$,

$$R(\omega)=\frac{NMp\tau K}{\epsilon(\omega)}\,\frac{d^2}{A_s},$$

where the symbols have the same meanings given above, and N represents the attenuation of a neutral-density filter used to switch the interferometer between gain-states 1 and 6. The uncertainty in M of 11 percent again dominates the errors. The $\Delta\tau/\tau$ is 3.7 percent at all wavelengths; $\Delta p/p$ varies from 2 to 6 percent; $\Delta\epsilon/\epsilon$ is about 2.5 percent throughout. Here the lamp is sufficiently far from the instrument (1 meter) that errors in d, as well as in A_s, are negligible. Thus, $\Delta R/R \approx 12$ percent for the silicon channel.

Table 24-2A shows the error contributions to $\Delta R/R$ for the lead selenide cell; Table 24-2B shows the error contributions for the silicon cell.

Table 24-2. Calibration Responsivity Error Analysis

A. Lead-Selenide Interferometer Channel

ω, cm^{-1}	$\Delta M/M$	$2\,\Delta d/d$	$\Delta p/p$	$\Delta A_s/A_s$	$\Delta\tau/\tau$	$\Delta\omega/\omega$	$[*]\cdot T\,\Delta\omega$	$[*]\cdot\omega\,\Delta T$	$\Delta R/R$
2400	±11%	±8%	±0.4%	±1.6%	±1.3%	±0.6%	±0.5%	±1.5%	±14%
4000	±11%	±8%	±0.5%	±1.6%	±1.9%	±0.4%	±0.5%	±2.3%	±14%
7000	±11%	±8%	±3.2%	±1.6%	±2.8%	±0.2%	±0.5%	±4.0%	±15%

$$[*] = \frac{1.4388 \exp\left(\frac{1.4388\,\omega}{T}\right)}{T^2\left[\exp\left(\frac{1.43888\,\omega}{T}\right) - 1\right]}$$

B. Silicon Interferometer Channel

ω, cm^{-1}	$\Delta M/M$	$\Delta N/N$	$\Delta p/p$	$\Delta\tau/\tau$	$\Delta\epsilon/\epsilon$	$\Delta R/R$
9500	±11%	±2%	±5.6%	±3.6%	±2.8%	±13%
10500	±11%	±2%	±2.2%	±3.7%	±2.6%	±12%
11500	±11%	±2%	±3.8%	±3.7%	±2.4%	±12%

24-5.5 Pre-flight Calibration

To verify that the instrument is in top working condition and has not been damaged during shipment, functional and detailed tests are performed as close to launch time as possible. These tests essentially duplicate the original laboratory calibration. In addition, a small field source on a mount is used. It has provisions for alignment with the optical axis of the instrument in the vehicle, and for variation of azimuth and elevation coordinates.

24-6 DATA LINK AND RECORDER INTERFACE

The scan rate of 10 scans/sec and the 20-cm^{-1} resolution for the Michelson interferometer precluded the use of a standard telemetry channel. Instead, it was necessary to use a specially designed X-band data-link system with four information channels to accommodate individually-modulated carriers. These channels are frequency-multiplexed across a bandwidth of 200 MHz centered near 8 GHz. The interferometer transmits data over analog data channels with 100-kHz bandwidth for each of the two detectors. The data received is recorded in real time and played back later in data processing. Chapter 25 contains a more detailed discussion of this system.

24-7 CONCLUSIONS

The flight performance of the instrument essentially matched the specifications given in Table 24-1. The flight experiments were successful. Typical examples of the interferograms received, and reduced spectra are shown in Figure 24-12 for the PbSe channel and in Figure 24-13 for the Si channel. Co-adding of interferograms has produced signal-to-noise ratios significantly improved from the individual-spectra values.

The Michelson interferometer-spectrometer has been utilized successfully as a primary infrared sensor

for space-vehicle applications. It has been shown that the Michelson interferometer can be made sufficiently rugged to provide high-quality spectral data despite a hostile acceleration, vibration, thermal, and vacuum environment.

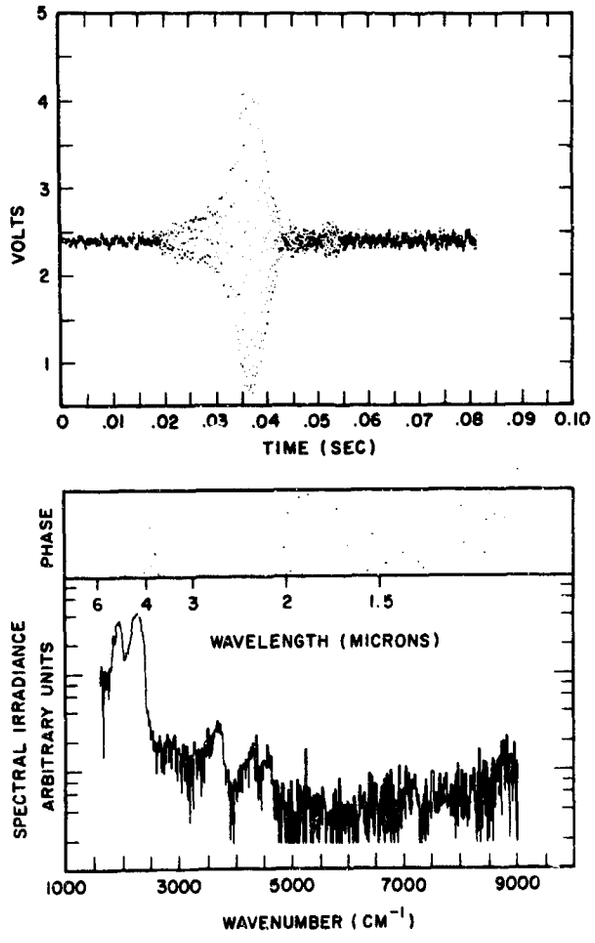


Figure 24-12. Typical Flight Interferogram for PbSe Channel with Corresponding Spectrum

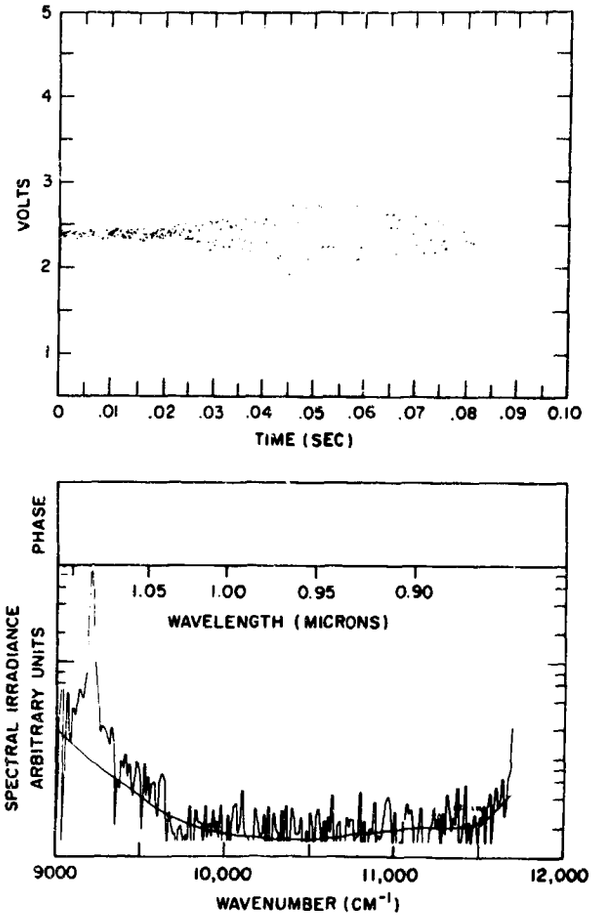


Figure 24-13. Typical Flight Interferogram for Si Channel with Corresponding Spectrum

Acknowledgments

Many persons in various laboratories, agencies, and independent companies have contributed significantly to the development of the flight interferometers and to the flight program that made possible their conception and application. Some of these contributors are C. D. Bass, S. Coleman, M. Emshwiller, L. C. Hebel, Jr., T. F. McIntosh, L. Mertz, G. R. Pruitt, L. D. Tice, and W. Van de Staat.

References

1. Fellgett, P. (1950) Johns Hopkins Infrared Seminar, quoted by Strong, J. (1957) *J. Opt. Soc. Am.* **47**: 354.

2. Gebbie, H. A., and Vanasse, G. A. (1956) *Nature* **178**: 432.

3. Connes, J. (1961) *Rev. Opt.* **40**.

4. Richards, P. L. (1964) *J. Opt. Soc. Am.* **54**: 1474.

5. Martin, A. E. (1966) *Infrared Instrumentation and Techniques*, Elsevier Publishing Co.

6. Mertz, L. (1965) *Transformations in Optics*, John Wiley.

7. Rayleigh, Strutt, J. W. (Lord Rayleigh) (1912) *Phil. Mag.* **24**: 864.

Discussion

Q. (D. P. Cruikshank, University of Arizona): What are you trying to observe with this instrument?

A. (Thomas P. Sheahen): Radiation from air species as the vehicle entered the atmosphere. The vehicle heated the air as it entered, thereby causing emission from CO_2, NO, etc.

Q. (Douglas McNutt, U.S. Naval Research Laboratory): How fast are your gain-switching amplifiers and is it feasible to switch gain during the recording of an interferogram?

A. (Thomas P. Sheahen): No. Gain will switch one step after any scan. It will not switch during an interferogram (there is a 0.02 dead time in between). Thus, it takes 7 scans (0.7 sec) to cover the full range of gain states.

Q. (Robert J. Chambers, University of California, Berkeley): What beamsplitter did you use for this rather wide spectral range instrument?

A. (Thomas P. Sheahen): Iron oxide. It wasn't the best, but it would go down to 8520 Å or lower, and wouldn't break.

Q. (Paul L. Richards, University of California): Is your method of chirping different in principle from the technique of inserting dispersive samples into one arm of an interferometer described earlier by E. Bell?

A. (T. Sheahen): Actually, I don't really know what Bell uses. Larry Mertz can answer this best.

Comment. (P. Richards): I believe that they are the same except for the use of a field compensation plate.

Comment. (L. Mertz, Smithsonian Astrophysical Observatory): Other exceptions are the motive and the degree of chirping. The chirping plates are not the samples being measured in the BTL experiment.

Q. What kind of vehicles were these interferometers flown on?

A. (T. Sheahen): A specially instrumented vehicle of about 1 m diameter carrying typically 10 different instruments and their telemetry, all designed to collect data below 100 km.

Contents

25. Data Recovery from a Michelson Interferometer-Spectrometer On-Board a Space Vehicle

Thomas P. Sheahen
Bell Telephone Laboratories
Whippany, New Jersey

Warren R. Howell
Block Engineering, Inc.
Cambridge, Massachusetts

Glenn F. Hohnstreiter
Bell Telephone Laboratories
Whippany, New Jersey

Isaiah Coleman
Block Engineering, Inc.
Cambridge, Massachusetts

Abstract

Interferometer measurements were telemetered to ground over an X-band data link. The reference interferogram determined the digitizing points on the data interferogram. Coadding of interferograms and phase-correction suppressed noise very effectively and enabled easy identification of important spectral features.

25–1 INTRODUCTION

In Chapter 24 (Michelson Interferometer-Spectrometer for On-Board Space Vehicle Measurements, by Hohnstreiter, Howell, and Sheahen), we described the instrument itself, from the design criteria through the optical system to the calibration and testing procedures. In this chapter we discuss the data transmission and processing, and the results of flight experiments.

The instrument output signal is first modulated onto a carrier signal, then further modulated onto a higher transmission frequency and telemetered to ground. Although the data link has a basic signal-to-noise ratio of only 30 dB, electronic and processing techniques result in spectra of 56-dB quality. The receiving station demodulates and tape-records the signals in real time. During later playback, the series of interferograms are digitized and prepared for computer handling. After a fast Fourier transform, the

resulting spectra are phase-corrected, plotted, and analyzed. By coadding interferograms it is possible to recover many spectral features lost in the noise of a single scan.

25-2 THE DATA LINK

Optical signals received by the interferometer are converted to output voltages between zero and 5 V. Figure 25-1 shows how these signals are then handled.

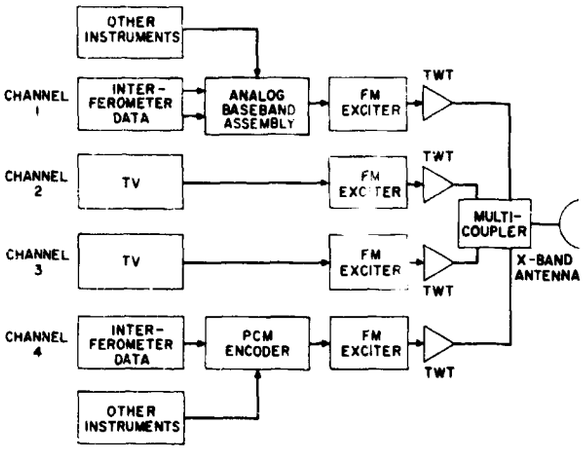


Figure 25-1. On-board Data Link Components

The output of each detector feeds one subchannel of an analog baseband assembly (ABA), which is the first link in the telemetry chain. This assembly consists of a frequency synthesizer that generates 15 subcarriers, 15 double-sideband amplitude modulators, and a summing amplifier system that multiplexes the resulting 15 subchannels into a 4-MHz bandwidth.

Although many instruments feed data to this assembly, the interferometer occupies two subchannels, each 100 kHz wide at baseband and centered at 853.2 and 1312.5 kHz. The long-wavelength lead selenide detector drives one of these; the short-wavelength silicon detector uses the other. Both output signals contain the reference signal superposed on the data interferogram.

In addition to the detector outputs, a variety of diagnostic measurements are also telemetered to ground. The interferometer outputs these measurements continuously, but because these measurements change slowly they are sampled only about 10 times a second (roughly once per scan) by a pulse-code modulation (PCM) system known as the digital data link assembly (DDLA). This device is parallel in function to the analog baseband assembly; it digitizes the voltages from the interferometer, multiplexes their signals (along with comparable signals from other instruments also in the package), and outputs a stream of 10^6 bits/sec.

Both the analog baseband assembly and the PCM system feed FM exciters, and two other 4-MHz video channels also feed such exciters. The exciter for the digital data uses frequency-shift keying (FSK) to represent 1's and 0's within a 7-MHz band centered at approximately 8 GHz (X band); the other three transmit their 4-HMz analog information by FM within 35-MHz bands carried at other frequencies, also near 8 GHz. The output of the exciters drives a set of traveling-wave tubes (TWTs) whose outputs are combined in a multicoupler and fed to the telemetry antenna. The four telemetered bands thus comprise a composite signal of approximately 200-MHz bandwidth centered at approximately 8 GHz for transmission through the atmosphere. Each of the four TWTs generates 35 W of power; but losses in the circuitry and waveguide preceding the antenna reduce the total radiated power to about 50 W. Ultimately this power suffices to guarantee virtually no errors in the received digital signals, and yields a basic carrier-to-noise ratio of 27 dB for the X-band analog carriers.

On the ground, the entire multiplexing procedure is inverted. The process is illustrated in Figure 25-2. Approximately 10^{-10} W are received by a 12-m-diam. dish antenna with a waveguide pickup at the focal point. The noise temperature of the sky plus antenna is 100° K. The received signal enters a cooled parametric amplifier, then a low-noise TWT, then another TWT with automatic gain control, and then a multicoupler, where the composite X-band signal is separated into its four individual X-band carriers. A set of 4 TWTs follow, after which each signal is limited and demodulated to provide the 4-MHz video channels. The effective operating temperature of the system receiver is 600° K. As in any FM system, the limiter effectively removes AM noise by clipping the received signals, leaving only the comparatively clean frequency information. After demodulation, the 27-dB carrier-to-noise ratio provides a 40-dB signal-to-noise ratio in the video channels. These signals feed into video amplifiers, and are recorded on wideband tape recorders.

Simultaneously, the analog baseband assembly channel is fed into a filter-detector, which restores the baseband signals to their original zero to 5-V condition seen at the instrument outputs. Unfortunately, this device degrades the signal-to-noise ratio to 30 dB for most of the analog subchannels. Redundant multiple-track tape recordings are made of the 15 subchannels to guard against sudden failures during the time that signals are being received. Two pairs of these redundant tracks contain the analog data output from the interferometer.

The digital signals are handled in a simpler way. The 'video' band of digital data is recorded directly on a 1.5-MHz bandwidth recorder, and the diagnostic voltages from the interferometer lie somewhere within the bit stream.

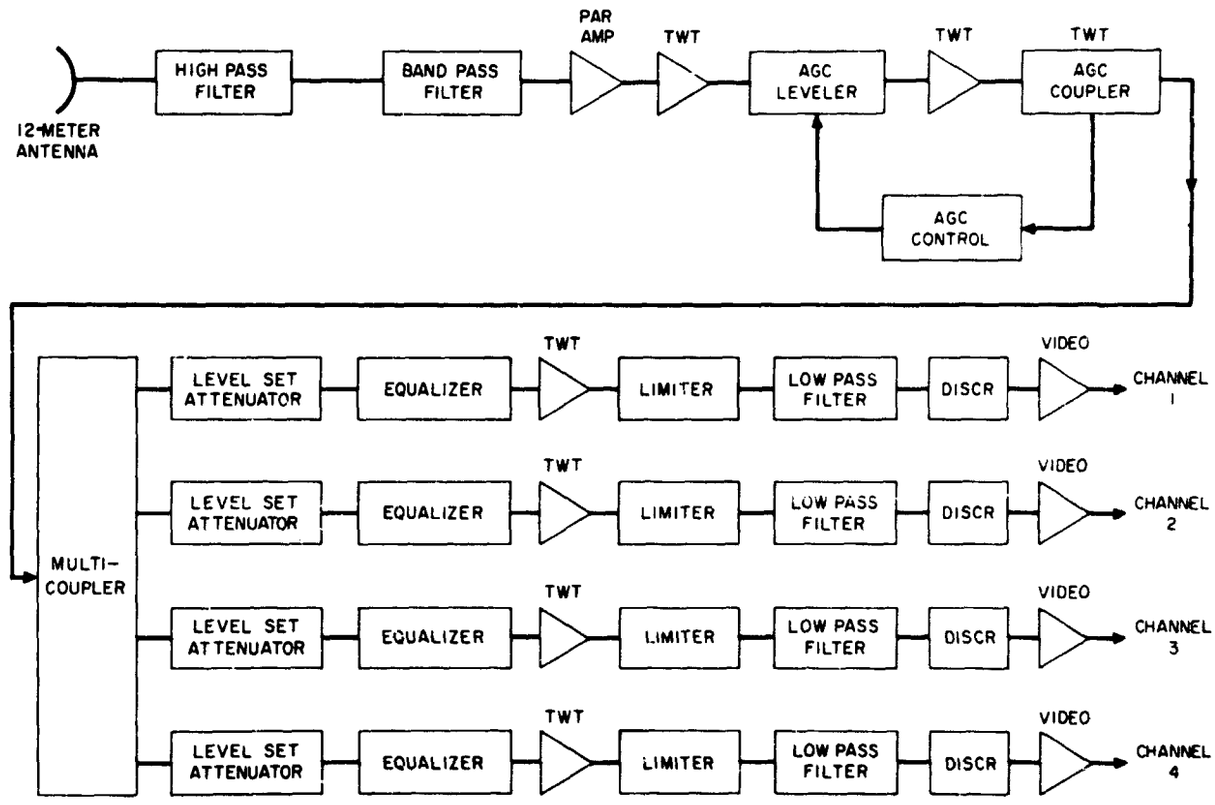


Figure 25-2. Ground Data Receiver System

All subsequent decomposition of signals is done later, at reduced tape-playback speeds.

There is one important difference between the interferometer subchannels and the analog subchannels of other instruments: the signal-to-noise ratio is 40 dB for interferometer data, instead of 30 dB as for other instruments. This is achieved by the following two complementary techniques.

First, filtering of the 100-kHz band reduces noise considerably. The reference signal can be separated by one narrowband filter centered at 25 kHz. Moreover, the data interferogram contains no frequency higher than about 15 kHz, corresponding to the short-wavelength cutoff of the optical filter and the mirror velocity, so it too can be filtered. Since there is no information above the first 15 kHz of the channel, nothing except noise is lost by discarding the higher frequencies, and the signal-to-noise ratio rises accordingly.

Second, preemphasis in the on-board analog baseband assembly gives extra power to the interferometer subchannels. Obviously, all this power appears in the lowest 30 kHz. As a result, after all detection, demodulation, and filtering, these two subchannels have signal-to-noise ratios of 40 dB.

25-2.1 Effect of Chirping

The need for a high signal-to-noise ratio was recognized early in the design stages of this experiment. Curve A of Figure 25-3 shows a synthetic spectrum typical of the expected radiation. Its Fourier transform defines an interferogram bearing no noise. Curves B, C, D, and E of Figure 25-3 show the successive degradation of the spectrum resulting from adding various levels of random noise to the interferogram. The rms noise level in Curve B is 0.001 of the peak signal, corresponding to a peak signal 60 dB above noise. A noise level of 0.003 of the peak signal (50 dB) produced Curve C; 0.01 (40 dB), Curve D; and 0.03 (30 dB), Curve E. Clearly, 60 dB is quite acceptable, 50 dB is marginal, 40 dB is unacceptable, and 30 dB is a catastrophe.

Since a chirped interferogram[1] has no pronounced central peak, the amplifier gain can be set higher without fear of saturating or clipping the signal. A noise level of 0.01 of this peak (peak signal 40 dB above the noise) is thus much less harmful. In Figure 25-4, Curve A shows the synthetic interferogram obtained by transforming and chirping our synthetic spectrum and then adding −40 dB of noise. Curve B is the spectrum that results from adding

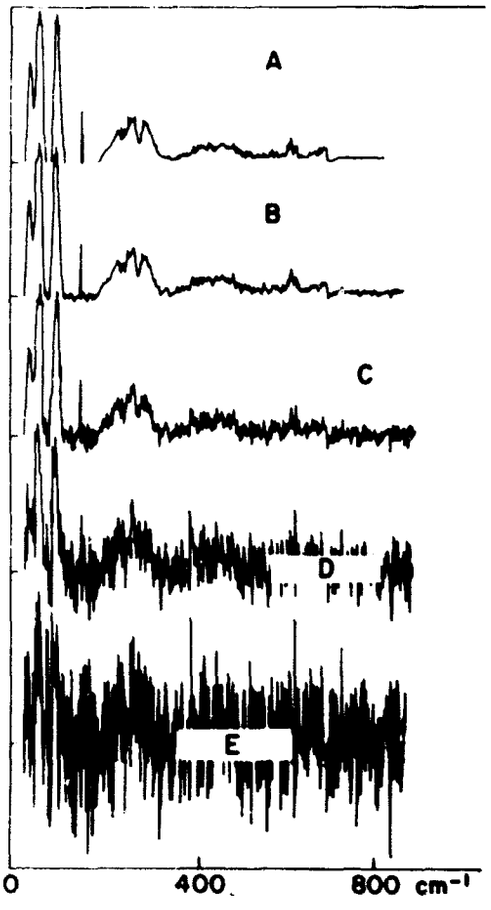


Figure 25-3. Unchirped Spectra [(*A*) synthetic spectrum of expected radiation; (*B*) spectrum obtained by adding 0.001 rms noise to interferogram from (*A*); (*C*) spectrum obtained by adding 0.003 rms noise to interferogram from (*A*); (*D*) spectrum obtained by adding 0.01 rms noise to interferogram from (*A*); (*E*) spectrum obtained by adding 0.03 rms noise to interferogram from (*A*)]

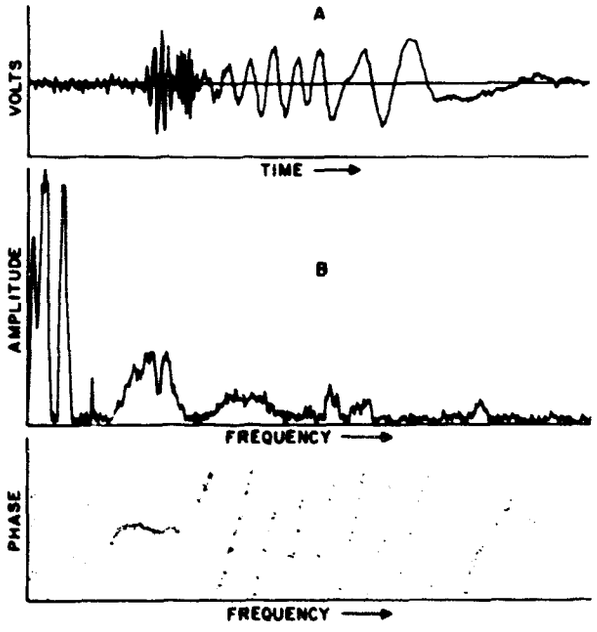


Figure 25-4. Effect of Chirping [(*A*) synthetic interferogram obtained by transforming and chirping synthetic spectrum (*A*) of Figure 25-3, with 0.01 rms noise added to interferogram; (*B*) spectrum obtained from interferogram 25-4(*A*)]

0.01 noise (40 dB) to the chirped interferogram. A comparison shows that this spectrum lies between Curves B and C of Figure 25-3. We have nominally called this signal-to-noise ratio 56 dB. Since the actual signal-to-noise ratio was only 40 dB, we credit the chirping process with 16 dB of gain. But it is essential here to avoid numerology; the important point is that a 40-dB chirped interferogram is just as good as a 56-dB conventional interferogram. The label '16 dB gain' for the chirping process means that the lower-contrast fringes of a chirped interferogram can tolerate 16 dB more noise than an unchirped interferogram, with no more degradation of spectral results.

It is by no means obvious that a different spectral source distribution would yield the same 'gain,' because the reduction in dynamic range is a function of the spectral distribution of the source. In fact, our synthesized projections were later verified by laboratory data. Figures 25-5 and 25-6 show laboratory spectra of a 1200° C glo-bar and an oxygen acetylene flame. Both spectra are from the Channel 1 lead selenide detector. For the blackbody spectrum, chirping reduced the interferogram dynamic range by 17 dB. A reduction of only 12 dB was obtained with the emission spectrum.

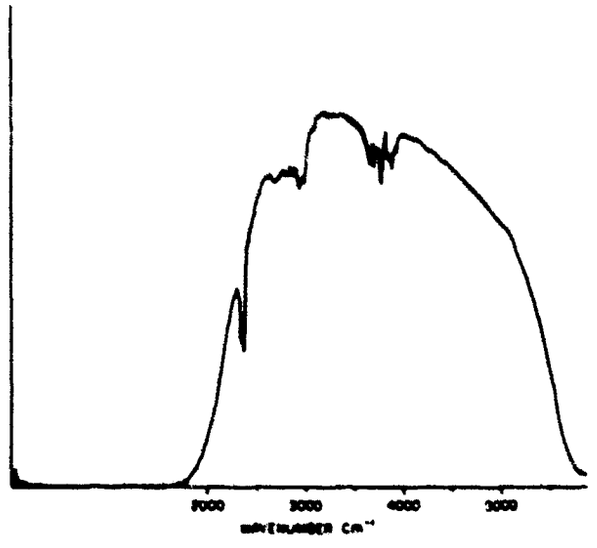


Figure 25-5. Laboratory Spectrum of a Blackbody Obtained With a Chirped Interferometer

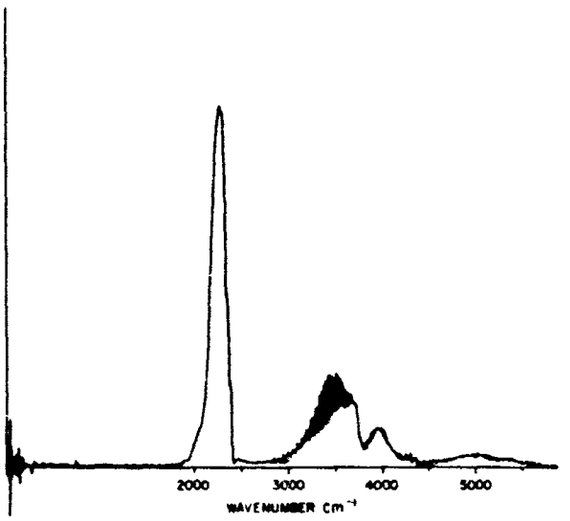


Figure 25-6. Laboratory Spectrum of Oxygen Acetylene Flame Obtained With a Chirped Interferometer

For our purposes, spectra of 36-dB quality [Figure 25-4 (*B*)] are good enough. By chirping, band-limiting, and preemphasis of the lower frequencies in the analog telemetry channels, we can obtain such spectra over a 30-dB signal-to-noise data link.

25-3 DIGITAL SAMPLING

The data tapes recorded in real time are played back at $\frac{1}{16}$th their original speed on a tape recorder with compensating filters. The digital bit stream is unpacked, sorted, and merged with other useful data, eventually resulting in a tape suitable for running on a conventional computing facility. The analog data follow a different path, and the interferometer channels receive special treatment.

Each 100-kHz bandwidth signal is outputted from a channel of the tape playback recorder. The signal is fed into a special demodulator that separates the reference interferogram from the data interferogram. The rolloff of these filter characteristics is steep, but not steep enough to prevent the reference signal from intermodulating the data signal, however, the purity of the data interferogram is restored in the next step. The two separate signals are fed to different points: the data interferogram (zero to 15 kHz) goes to a conventional analog-digital converter, the reference interferogram goes to a network that detects the positive zero crossings and sends a command to the data A-D converter upon each such crossing. Thus, the data interferogram is sampled only at those moments that the reference signal is zero; hence, the intermodulation contribution is likewise zero, and the true value of the interferogram is obtained.

Despite the servomechanism that corrects the mirror motion, irregularities in the mirror drive and tape recorder flutter can distort the data interferogram. The reference signal compensates for all of these. The digital representation of the data interferogram is uniformly spaced in optical retardation. Normally, one scan is represented by 2048 points; if a short scan occurs, the additional points required are filled in with zeroes. Such an artifact corresponds to a peculiar apodizing profile, but the numerical effect on the final spectrum is negligible. In most cases the number of points sampled was sufficient to extend well into the wings of the chirped interferograms.

At the start of each scan, the instrument injects a starting pulse on the reference signal to enable the digitizing equipment to define the beginning of a new interferogram. The pulse is derived from the white light fringe of the incandescent lamp in the reference cube. Details of the circuitry and fringes occasionally make this pulse one cycle late; more important, the digitizing equipment may lose a cycle by detecting the end of the pulse rather than its beginning. Although these random events are infrequent, they are ultimately responsible for imperfect coherence when coadding interferograms.

25-4 BULK PROCESSING

The data from each experiment consisted of two tapes: one for the long-wavelength detector and one for the short-wavelength detector.

The interferograms were apodized before Fourier-transforming. A variety of apodizations were used but there was very little difference between any of the resulting spectra, primarily because there were few sharp lines in the spectra. For most of the processing, a simple boxcar apodization was used, blanking out 100 points on each extreme of the data and leaving the rest unchanged. The only instance in which this caused any difficulty was in the silicon channel when the on-board calibration lamp dominated the radiation from air species. Then the $(\sin x)/x$ transform of the boxcar-apodization profile raised sidelobes of the calibration lamp to the point of obscuring surrounding data. A trapezoidal apodization profile that was contrived to compensate for some "average" amount of chirp did reduce sidelobes but produced no other noticeable effects. It appears that the distortions introduced by alternative apodization methods are negligible compared with the effects of noise on the interferogram.

The raw spectra were calculated from the apodized interferograms by means of the Cooley-Tukey[2] algorithm. The long-wavelength-band data were read into the real part of a complex array, and the short-wavelength-band data became the imaginary part. The Fourier transform

$$S_w = \sum_{i=0}^{N-1} g_i \exp(-2\pi i w i / N)$$

was then calculated.* After this the two separate spectra were obtained by orthogonalizing the real and imaginary transforms (W. T. Hartwell, private communication). This gave:

$$S_a(\omega) = \tfrac{1}{2}[S(\omega)+S^*(-\omega)]$$

$$S_b(\omega) = \tfrac{1}{2}[S(-\omega)-S^*(\omega)].$$

By comparison with separate FFTs for each channel, this route saved computer time despite the small added steps of orthogonalization. Each resulting complex spectrum was then represented by an amplitude and a phase:

$$S_a(\omega) = A_a(\omega) \exp\{i\phi_a(\omega)\}$$

$$S_b(\omega) = A_b(\omega) \exp\{i\phi_b(\omega)\}.$$

At this point the amplitude spectrum is known, and the optical intensities can be outputed via printer, plotter, or tape. After multiplying by a calibration factor, the power spectrum $A(\omega)$ thus obtained is sufficiently accurate for our purposes, but all the spectra were phase-corrected in order to improve the signal-to-noise ratio.

In an unchirped interferogram, the signal is supposedly the cosine transform of the spectrum, so (assuming precise sampling) the phase can be anticipated to be zero. Mertz[1] has shown how phase-correction can relax this zero phase requirement to a mere stationary phase, by eliminating the sine transform of the interferogram. For a chirped interferogram, the optical properties of the chirping medium as well as the phase spectrum must be known. Computationally, a ninth-order polynomial specifies the anticipated phase shift accurately at every wavelength of interest.

Whenever noise is superimposed on an interferogram, the Fourier transform of the noise is mixed in with the true spectrum producing the interferogram. Since the Fourier transform is a linear operator, the complex Fourier transform $S(\omega)$ is the linear superposition of the true spectrum and the noise spectrum:

$$S(\omega) = A_t(\omega) \exp\{i\phi_t(\omega)\} + N(\omega) \exp\{i\phi_n(\omega)\},$$

where $A_t(\omega)$ is the true spectrum, $\phi_t(\omega)$ is the true phase shift, and $N(\omega)$ and $\phi_n(\omega)$ are the amplitude and phase of the noise spectrum. Trouble occurs when the nonlinear operation of taking the square modulus inextricably mixes the signal and noise:

$$|S(\omega)|^2 = A^2(\omega)+N^2(\omega)+2A(\omega)N(\omega)\cos(\phi_t-\phi_n). \tag{25-1}$$

This familiar formula displays the established fact that signal and noise add vectorially in the complex plane. The projection of $S(\omega)$ along $A(\omega)$ is therefore

$$S(\omega)\cos(\phi_s-\phi_t) = A(\omega)+N(\omega)\cos(\phi_t-\phi_n), \tag{25-2}$$

where $\phi_s(\omega)$ is the measured phase angle.

It is easy to show that the out-of-phase component of noise is

$$N(\omega)\sin(\phi_t-\phi_n) = S(\omega)\sin(\phi_s-\phi_t),$$

in which case solving a quadratic leads to

$$A(\omega) = S(\omega)[\cos(\phi_s-\phi_t) \pm \sin(\phi_s-\phi_t)\cot(\phi_t-\phi_n)].$$

The phase of the noise is unknown, however, so the only correction that can be made is the simple projection of $S(\omega)$ upon $A(\omega)$ given by Eq. (25-2). Since noise is assumed to have random phase, this correction reduces noise by a factor of $1/\sqrt{2}$, which corresponds to a 3-dB improvement in signal-to-noise ratio.

One very important experimental artifact complicates the process of phase-correction. While the instrument is acquiring data in its dynamic environment, thermal and mechanical changes in the interferometer cube introduce slight changes in the location of the points at which the interferogram is sampled. This requires correction of the constant and linear terms[1] in the polynomial for each scan. The corrections are derived from spectral regions of strong signal-to-noise ratio, for which the assumption $\phi_s \approx \phi_t$ is locally justified. These empirical 'true' values of phase are used to update the phase polynomial by an iterative method, which adjusts all the phases in the high signal-to-noise regions, recomputes the residual

* Several extremely fast routines written in FORTRAN IV, especially FFT842 used here, were supplied by G. D. Bergland of Bell Telephone Laboratories, Whippany, N.J. Any inquiries should be addressed to G. D. Bergland.

phase differences, adjusts again, and continues until the cumulative residuals fall below a prespecified limit. This scheme adds about 0.1 min to the processing time of each spectrum.

The computer time for reading, apodizing, transforming, orthogonalizing, phase-correcting, plotting, and printing the data from 300 scans of 2048 points each in both channels amounted to 9 min of IBM 360/85 time. Of this, the fast Fourier transform subroutine required 60 msec per scan.

25–5 ERROR ANALYSIS

The calibration procedures and their associated errors are discussed in Chapter 24. Here we present a discussion of the total errors present in the flight experiment.

The object of the measurement is to find the spectral irradiance $H(\omega)$ of the radiating species. This is related to the telemetered voltages as

$$H(\omega) = \frac{\mathfrak{I}\lambda V(t) G_n}{R(\omega)},$$

where $\mathfrak{I}$ denotes the Fourier transform operator; λ is the normalizing constant that defines the relative signal level (when $\lambda = 1$, 5 V telemetered equals 5 V at the computer); $V(t)$ is the signal received over the telemetry link; G_n is the ratio of amplifier gain in gain state n to the amplifier gain in gain state 1; and $R(\omega)$ is the responsivity of the instrument. Errors in $R(\omega)$ comprise the totality of errors in calibration.

Assuming all errors are random and nonsystematic, we would obtain

$$\left(\frac{\Delta H}{H}\right)^2 = \left(\frac{\Delta V}{V}\right)^2 + \left(\frac{\Delta \lambda}{\lambda}\right)^2 + \left(\frac{\Delta G}{G}\right)^2 + \left(\frac{\Delta R}{R}\right)^2.$$

Unfortunately, this simple vector addition formula obscures the fact that voltage (and errors in voltage) are functions of time, whereas the error ΔH in spectral irradiance varies with frequency.

In considering the error in voltage received, the temptation is to trivially cite the 40-dB signal-to-noise ratio of the telemetry link and therefore set $(\Delta V/V) = 0.01$. To do so would be misleading, for two major reasons.

First, the noise in an FM transmission system is a percentage of full scale whether the full dynamic range is used or not. Since the noise level ΔV is nearly constant in absolute magnitude, rather than in percentage, the representation $(\Delta V/V)$ is pointless. Noise affects every fringe of the interferogram equally, regardless of the magnitude of the fringe. Thus, an interferogram produced by a clustered spectrum (whose central peak is spread very little by chirping) will accumulate a substantial noise contribution from the wings. By contrast, a spectrum of sustained high-contrast fringes will be comparatively immune to noise, with correspondingly less degradation of the final spectrum. Clearly, the error contribution by voltage errors depends on both the size and shape of the interferogram.

Second, the noise is random and has a flat spectrum whereas the signal is structured with various lines and continua. The noise is therefore redistributed 'unfairly' throughout the spectrum. For example, consider a cluster of lines whose intensity is an order of magnitude below a strong continuum at another frequency. The interferogram will be dominated by the stronger source, which will determine the gain state of the amplifier. The weaker source will ride along as modulation on the interferogram, out of the noise by a factor of only 10 instead of 100 like the strong signal. The continuum will be recognized in spite of its accompanying noise, but the line structure will be partially lost in the noise. Because the stronger signal triggers a change to a less sensitive gain state, the weaker lines suffer a disadvantage that would not occur in the absence of the strong source.

Recognizing these anomalies, we are forced to quote separate figures for the contribution of telemetry noise to the error in spectral irradiance of each channel at various wavelengths, intensities, and spectral distributions. This difficulty clutters all other discussion of errors in the final spectral irradiances. For now it suffices to state that the range of such error is between 2 and 36 percent in all cases.

The signal level parameter λ is a much less sophisticated error source in the data acquisition system. Because it is independent of time and frequency, it passes through the Fourier transform and uniformly affects all values of $H(\omega)$. The raw analog data tape of early units contains no fiducial 5-V signal to ensure precise setting of the digitizing level; this signal was supplied by later interferometers. The reference signal is nominally a 1-V peak-to-peak square wave, but a succession of filters and playback recorders degrade its shape prior to the digitizer. Therefore, in attempts to use this signal to define a fiducial 1-V digitizing level, the difficulty of distinguishing its peaks and valleys introduced error. Statistical analysis of the digital records showed fluctuations in λ of ± 17 percent. On later instruments, this deficiency was corrected by including a 5-V signal generated by the interferometer; this reduced error in λ to only ± 5 percent.

The error contribution of the successive gain states G_n of the amplifier is due to uncertainty in the attenuation introduced by switching the gain. The attentuation factor is used to convert the actual telemetered volts to equivalent gain state 1 voltage. Since the calibration was performed with the instrument in gain state 1, the uncertainty in gain for that state is absorbed into uncertainty in $R(\omega)$. The attenuation

factors were measured several times with a reproducibility of ±10 percent.

Some of the more attenuated gain states were 'afflicted' with a source of error more severe in principle. The detectors were not perfectly linear at very high intensity, which had the effect of making the gain factor G_n dependent on voltage, thereby distorting the true signal. Distortion can be represented as higher harmonics of a fundamental frequency; indeed, some minor spectral features observed at double and triple the frequency of strong sources can only be explained in terms of a few percent distortion at high intensity. [The problem of recovering signals passed through nonlinear amplifiers and transmitted over band-limited systems has been treated by Landau.[3]] In analyzing the data, we considered the effects of nonlinearity to be less than 5 percent and therefore minor compared with other error sources. Accordingly, nonlinear recovery techniques have not been used here.

The error in responsivity calibration $\Delta R/R$, taken from Table 24-2 (Chapter 24), was dominated by inaccuracy in maximum spectral volts of the calibration spectrum $\Delta M/M$. It is now possible to explain the origin of this error source. Recall that during calibration, signals generated by the interferogram pass through the data reduction system. To eliminate noise, about 100 of these interferograms are coadded and transformed to produce the calibration spectrum. Statistical studies of a large number of such coadding experiments have shown that net amplitude values deviate from the maximum of a single-scan spectrum between zero and −10 percent. A scatter in the start of the interferogram (defined by detecting the trigger pulse) of only 1 cycle can easily account for this fluctuation, when we consider the effect of coadding signals slightly out of phase. Several techniques are used to minimize this error, but it has not been entirely eliminated. Moreover, the accuracy of the signal level upon playback during calibration is upset by the same conditions that introduce error in λ, so the error in volts-to-counts conversion is again ±5 percent, as in flight. These two contributions sum vectorially to 11 percent error in maximum spectral volts M.

Table 25-1 summarizes the errors in the final values of $H(\omega)$ for four typical cases: lead selenide at high and low levels, and silicon at high and low levels. Inspection of this table suggests 30 percent as an appropriate number for the overall accuracy of the measurements. Without the telemetry noise contribution ΔV, the silicon cell would generally have an error of 16 percent and the lead selenide cell an error of 19 percent.

Calibrations performed at two independent laboratories showed typical agreements of 10 to 30 percent. One instrument showed a discrepancy of 50 percent, tentatively ascribed to incoherent coadding.

Table 25-1. Interferometer Error Summary

ω (cm^{-1})	$\Delta V/V$ Noise Error (%)	$\Delta\lambda/\lambda$ Signal Level, Early Version (%)	$\Delta\lambda/\lambda$ Signal Level, Later Version (%)	$\Delta G/G$ (%)	$\Delta R/R$ (%)	$\Delta H/H$, Early Version (%)	$\Delta H/H$, Later Version (%)
PbSe, LOW-LEVEL INTERFEROGRAMS							
2200	±20	±17	5	±10	±15	±32	±27
3500	± 5	±27	5	±10	±15	±25	±19
PbSe, HIGH-LEVEL INTERFEROGRAMS							
2200	± 7	±17	5	±10	±15	±26	±20
3500	± 2	±15	5	±10	±15	±25	±19
Si, LOW-LEVEL INTERFEROGRAMS							
9050	±36	±17	5	±10	±12	±43	±39
10900	±25	±17	5	±10	±12	±34	±30
Si, HIGH-LEVEL INTERFEROGRAMS							
9050	±28	±17	5	±10	±12	±36	±32
10900	± 6	±17	5	±10	±12	±24	±17

A large class of other errors have not been discussed, but they are all below 5 percent and contribute negligibly to errors in the spectra. Among these are tape recording and A–D conversion errors, frequency effects in electronic components, systemic errors in irradiance standards and measuring equipment, interpolations in calibration curves, and tilt of the optical axes.

25–6 RESULTS

During several suborbital flights, instruments of this type collected data on air species radiating in the range 0.85 to 5.3 μm. Figure 25-7 (top) shows a

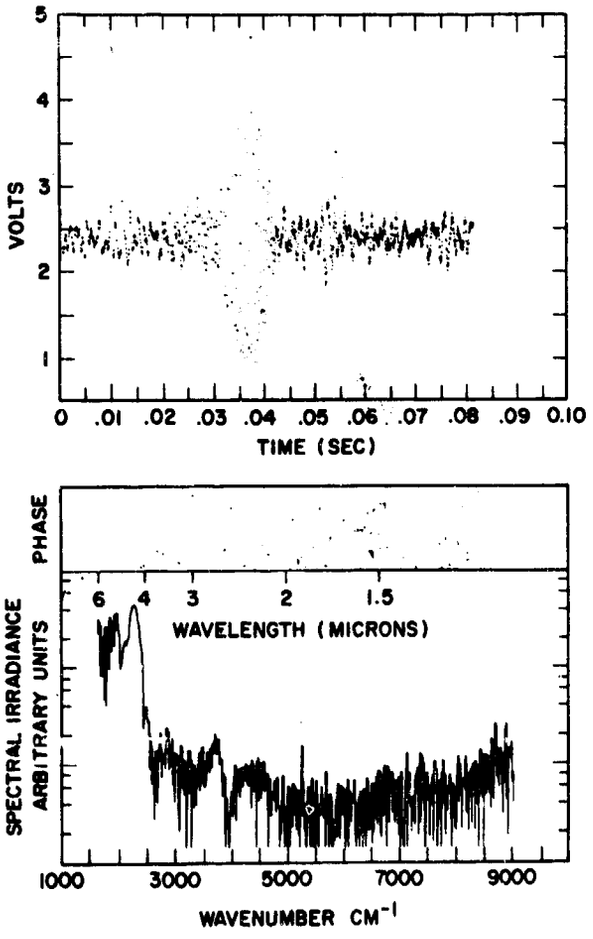


Figure 25-7. Flight Interferogram (top) and Spectrum (bottom) Obtained From the Lead Selenide Channel (severe vibration)

chirped interferogram obtained by the lead selenide detector during severe vibration in flight. Because the vibration frequencies were all low, the devastating

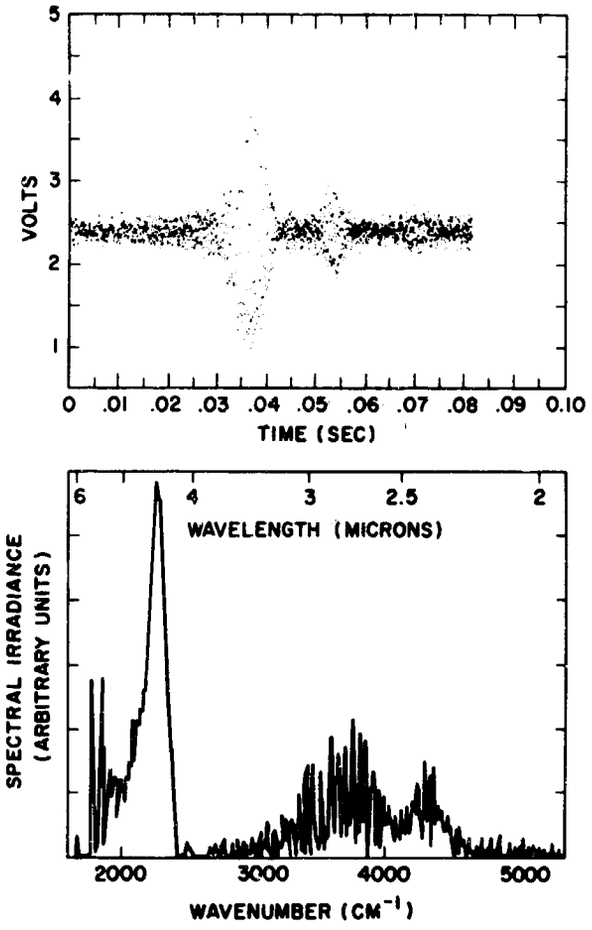


Figure 25-8. Flight Interferogram (top) and Spectrum (bottom) Obtained From the Lead Selenide Channel (typical)

effect on the interferogram was discarded from the spectrum, and vibrational noise was mitigated. A more typical interferogram from the lead selenide channel appears in Figure 25-8 (top) with its spectrum (bottom). Figure 25-9 (top) is a typical silicon chirped interferogram with its spectrum (bottom). Figure 25-10 is a rather weak spectrum from this channel.

Obviously, the effects of noise detract substantially from the quality of the spectra. To suppress noise, we have coadded a number of interferograms. The coherent signals grow as the number N of interferograms coadded, but the noise grows only as $\sqrt{N}$, which gives a net suppression of noise as $1/\sqrt{N}$. Coadding 20 scans gives very adequate quality in the resulting spectra.

Figure 25-11 is the spectrum resulting from coadding the 18 interferograms near Figure 25-7; Figure 25-12 shows the 25-scan average near Figure 25-8; Figure 25-13 was made by coadding 25 interferograms near Figure 25-9; Figure 25-14 is the 20-scan average near Figure 25-10.

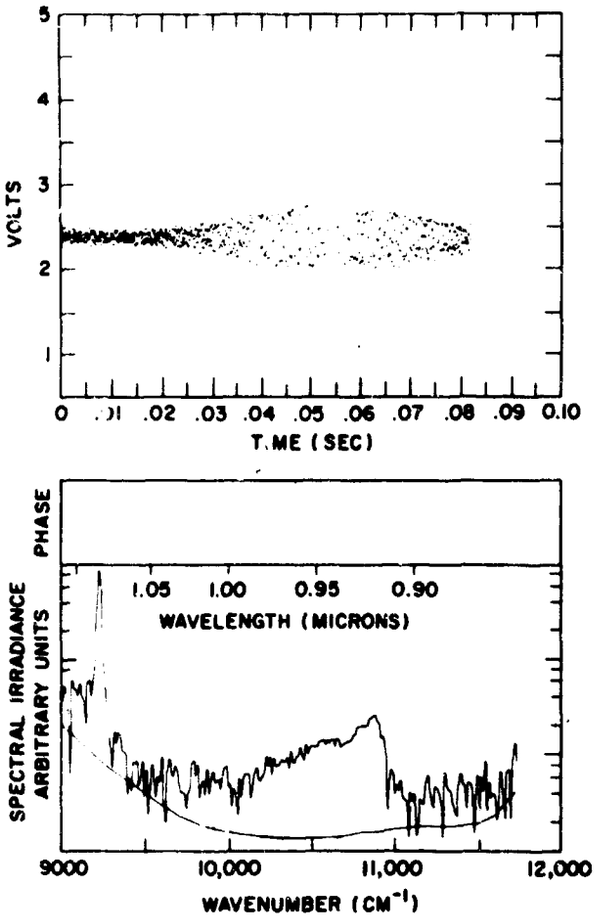


Figure 25-9. Flight Interferogram (top) and Spectrum (bottom) Obtained From the Silicon Channel (typical)

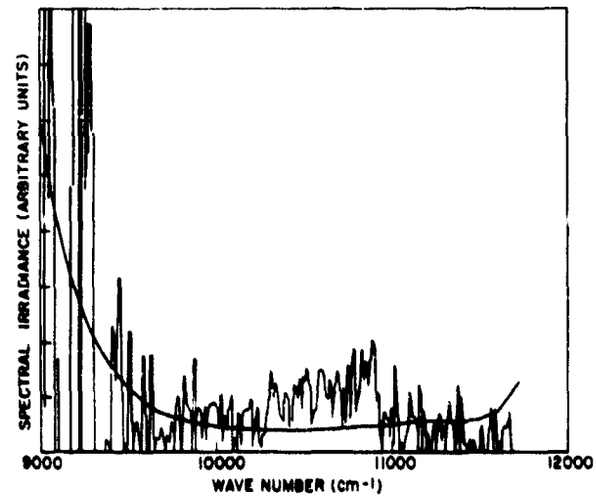


Figure 25-10. Flight Spectrum Obtained From the Silicon Channel (weak)

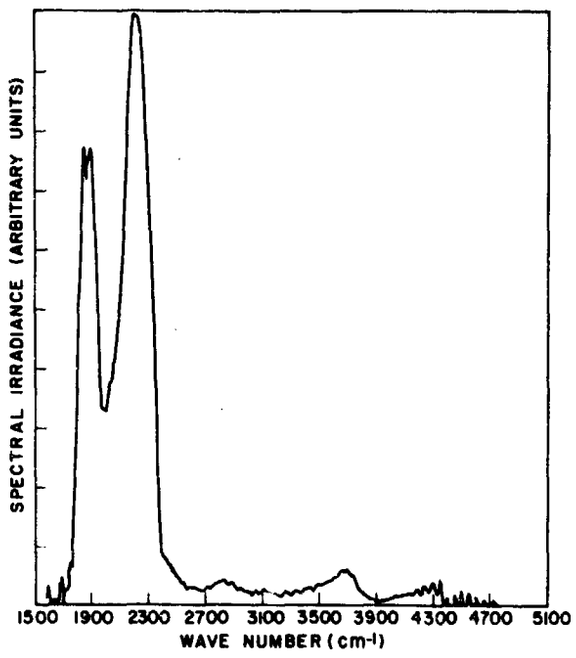


Figure 25-11. Coadded Spectrum Obtained From the Lead Selenide Channel (conditions as in Figure 25-7; 18 interferograms)

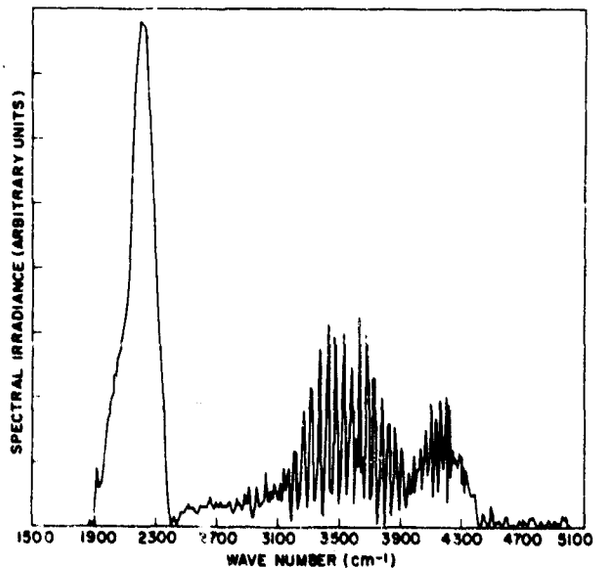


Figure 25-12. Coadded Spectrum Obtained From the Lead Selenide Channel (conditions as in Figure 25-8; 25 interferograms)

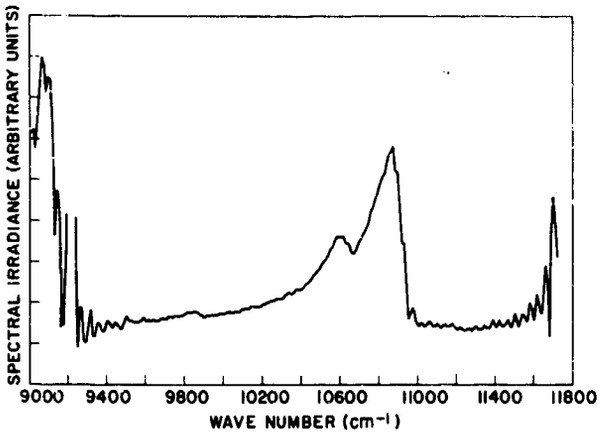


Figure 25-13. Coadded Spectrum Obtained From the Silicon Channel (conditions as in Figure 25-9; 25 interferograms)

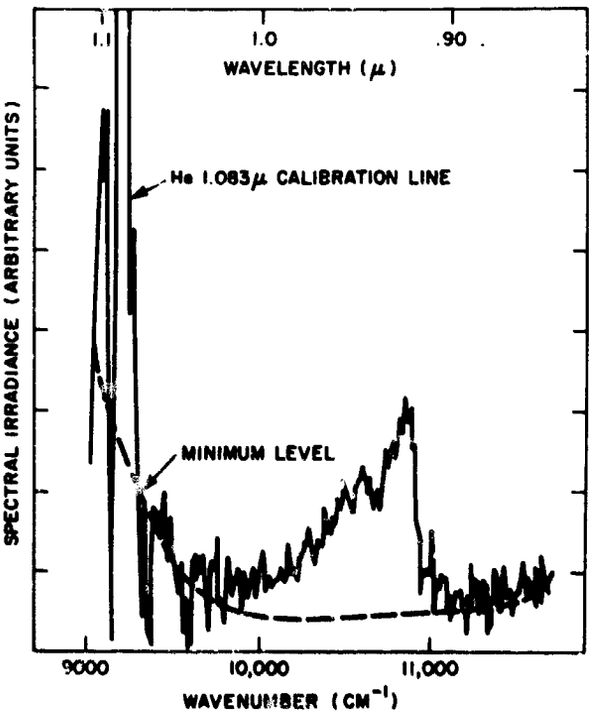


Figure 25-14. Coadded Spectrum Obtained From the Silicon Channel (conditions as in Figure 25-10; 20 interferograms)

A comparison of Figures 25-7 and 25-11 shows that spectral features only hinted at in the single scan stand out in the coadded version. Figures 25-8 and 25-12 demonstrate how coadding cleans up continua and establishes the validity of structural features lost in the noise of a single scan. Figures 25-9 and 25-13 contrast the effect of the coadding process on molecular radiation bands: the clean spectra of Figure 25-13 enable us to calculate a temperature, whereas those of Figure 25-9 do not allow unambiguous temperature determinations. Figure 25-10, a low-level signal, has few recognizable features beyond the calibration lamp, but Figure 25-14 contains clearly identifiable spectra.

For the higher frequencies, coadding is useful only when all the interferograms are precisely synchronized at their starting points. Frequencies close to the Nyquist frequency cannot tolerate even one fringe displacement in synchronism without losing their coherence; and once coherence is lost, there is no gain over the incoherent noise in coadding. Lower frequencies are correspondingly less sensitive to such asynchronism.

Synchronizing errors arise mainly from failure of the digitizing apparatus to associate the initial spike on the reference signal with one unique 'first' point in the interferogram. If the sampling circuitry delays one zero-crossing while discerning the initializing command, the interferogram will be out of phase by one reference fringe. Such an amount is not harmful to the low frequencies that dominate the lead selenide channel but it sacrifices coherence in the silicon channel. Fortunately, this phenomenon occurs relatively rarely, and is caused primarily by telemetry noise when it does occur.

We show no comparisons of the results of various apodizing functions. The only difference we found was in the suppression of sidelobes around the intense peak of the calibration lamp in the silicon channel.

More to the point for further study is the removal of false harmonics by computational techniques.[3] When significant low-level spectra occur at harmonics of intense lower-frequency sources, the nonlinear detector characteristic creates higher harmonics that obscure the naturally occurring low-level spectra. The removal of such artifacts will allow a better understanding of the observed spectra in the higher frequency ranges.

25-7 CONCLUSIONS

These results are only samples of a much broader collection of data whose interpretation has not yet been published. They illustrate the type of data obtainable with a Michelson interferometer applied to on-board measurements from a space vehicle.

The interferometer-spectrometer has been found to function properly in flight and to deliver information of high quality. The chief restraint on using the interferometer in this way is the need for a very sophisticated analog data link and ancillary ground equipment. An interferometer accompanied by a fast on-board analog-to-digital converter (triggered by the reference interferogram) would be relatively free of the signal-to-noise problems, signal-level errors, playback distortion, and related obstructions that augment the error in our measurements. Many space experiments have shown digital telemetry to be essentially noise-free over vast distances, and the application of such techniques to space interferometry would be a welcome advance.

Acknowledgments

Our author list should read like a synchrotron paper: many individuals have contributed to the total project, which is much bigger than the interferometer portion. In particular we wish to point out that the data link is attributable to R. E. Markle, T. F. McIntosh, and N. B. Rowe of Bell Telephone Laboratories; chirping and phase-correction originated with L. H. Mertz while he was with Block Associates; and the error analysis is drawn from the work of E. Tagliaferri of TRW systems. We have also benefited greatly from discussions with L. D. Tice of the Safeguard System Command, L. C. Hebel of Sandia, and C. D. Bass of Aerospace Corporation.

References

1. Mertz, L. (1965) *Transformations in Optics*, Wiley.

2. Cooley, J. W., and Tukey, J. W. (1965) *Mathematics of Computation* **19**: 297.

3. Landau, H. J. (1960) *Bell System Tech. J.* **39**: 351.

Discussion

Comments (Thomas Walter, AFCRL): In checking the ability to coadd interferograms, especially when there is doubt concerning the position of the white light reference, I have found it quite helpful to perform a cross correlation between the interferograms. This will determine if the interferograms must be shifted before coadding. Correlation can also be used to determine the skipping of a reference fringe in an interferogram.

Q. What were the molecular bands in your results?

A. (T. P. Sheahen): Radiation from air species excited by friction with the moving vehicle.

Q. What type of vehicles were the instruments flown on?

A. (T. P. Sheahen): A suborbital-trajectory object designed for upper atmospheric measurements under a severe environment.

Q. Meteorites?

A. (T. P. Sheahen): No.

Q. What was the spectrum appearing in Figures 25-8 and 25-12?

A. (T. P. Sheahen): Hydrogen fluoride. It is a product of ablation from the vehicle's surface.

Contents

26. Computer-controlled Telescope and Interferometer for Eclipse Observations

William G. Mankin, Robert M. MacQueen, and Robert H. Lee
High Altitude Observatory
National Center for Atmospheric Research*
Boulder, Colorado

Abstract

A two-beam telescope and interferometer for eclipse observations in the 8- to 13-μm range is described. A small general-purpose digital computer is an integral part of the experimental system, used for both equipment control and data analysis. The limit of radiometric sensitivity is set by fluctuations in sky emission.

26–1 INTRODUCTION

The two-beam telescope and interferometer that we have constructed for observing thermal emission from the F corona in the 10-μm region incorporates several unusual features, the most interesting of which is a small general-purpose digital computer used as an integral part of the experimental system rather than as a device for *post facto* data analysis. The computer controls the interferometer and collects the data as well as computing the Fourier transform of the interferogram. The use of the computer provides great flexibility; many observations require only a change of program rather than of hardware.

The instrument is designed for field work. It has been successfully used in a remote region of southern Mexico under conditions of high humidity, varying temperatures, and dust. In some respects it is more complicated than was required for the experiment performed, but it has served the additional purpose of testing the concepts and the hardware for planned observing programs, including airborne observations, where the more complex instrumentation is needed.

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

26–2 INSTRUMENTATION AND TECHNIQUES

Figure 26-1 is a diagram of the instrument optics. An f/8 Cassegrain telescope of 15-cm aperture covers two fields of view, each approximately 5 min of arc in diameter, separated in space by 1.5°. One field sees the corona; the other looks at the sky 1.5° away. By taking the spectrum of the difference we eliminate the sky emission, which is expected to be much larger than the emission from the corona.

The radiation from each field of view is collimated into a 2-cm-diam beam. A reflecting chopper alternates the two beams onto the entrance of a scanning Michelson interferometer. After passing through the interferometer, the radiation is focused onto a gallium-doped germanium bolometer cooled by liquid helium pumped to a temperature of 2° K. The dewar has a capacity of 1.33 l of liquid helium and a hold time of 9 hr after reaching the operating temperature. A low-noise FET preamplifier is mounted directly on the dewar. At the chopping frequency of 72 Hz, the rms noise equivalent power is approximately 1×10^{-14} W Hz$^{\frac{1}{2}}$. The window and internal lens are made of barium fluoride. An interference filter, cooled to liquid helium temperature, limits the

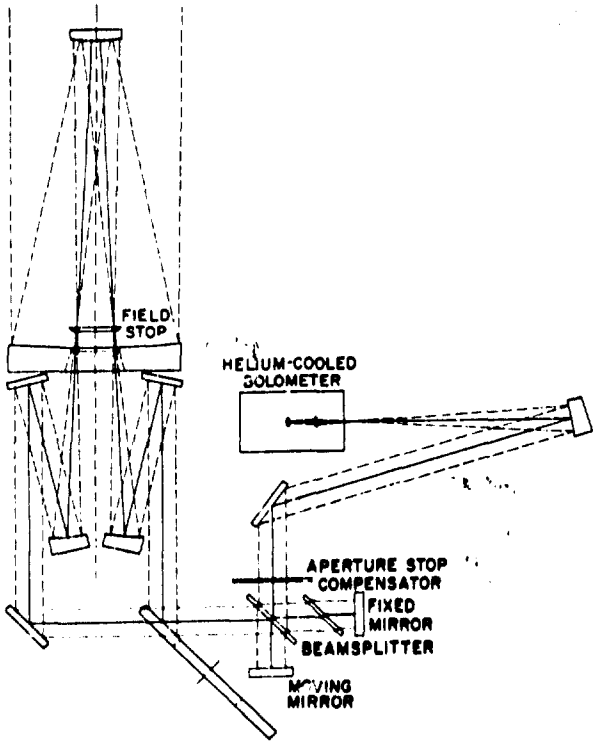


Figure 26-1. Diagram of Two-beam Infrared Radiometer and Interferometer Optics

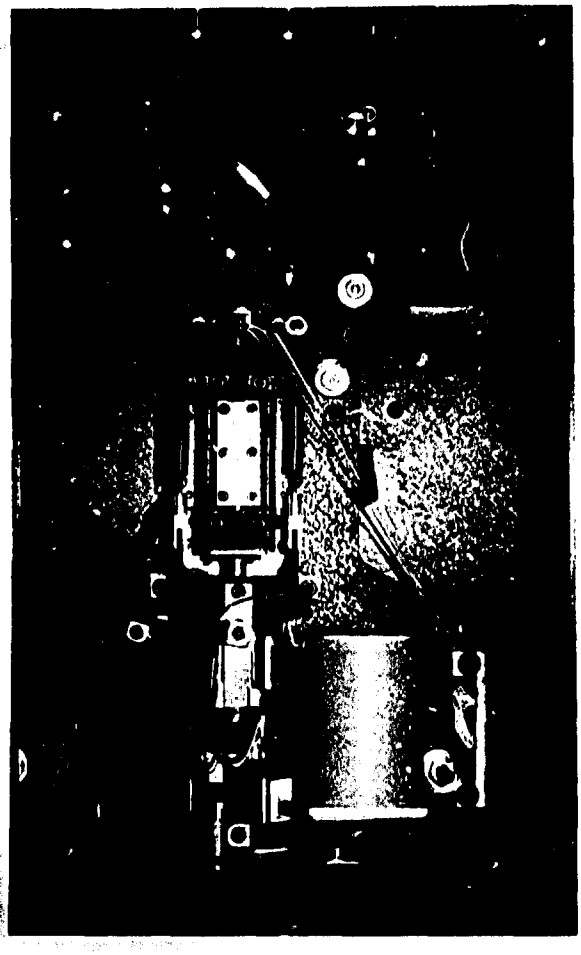


Figure 26-2. Scanning Michelson Interferometer for Coronal Observations

spectral bandpass of the system to 7 to 13 μm, and reduces the background radiation on the detector.

The interferometer is shown in Figure 26-2. The beam splitter is germanium, 2 mm thick, with an antireflection coating on one side and a partially reflecting coating on the other; its efficiency, compared with that of an ideal beam splitter, is greater than 90 percent over the 8- to 12-μm range, averaged over both polarizations. The compensator is also germanium and is antireflection-coated on both sides. A difference in thickness of 75 μm between the beam splitter and compensator produces a small amount of chirping, but this is not troublesome.

The moving mirror is mounted on a shaft supported in two precision ball bushings. A stepping motor turns a differential screw micrometer translator to effect the mirror motion. The mirror travel is 1 mm, with zero path difference at the center of the travel; the resulting resolution is 5 cm^{-1}.

The entire instrument mounted in its two-axis gimbal system is shown in Figure 26-3. The two integrating gyros each have a single degree of freedom. When they sense any rotation of the instrument in inertial space they produce an error signal, used to drive torque motors on each axis to correct the error. In this way a pointing accuracy of better than 10 sec of arc is achieved. There are provisions for driving the instrument about either axis at a slow or a fast rate to permit spatial scans.

Figure 26-4 is a block diagram of the complete observing system. The output of the detector is amplified and digitized. As a backup system it also goes to a standard lock-in amplifier and strip chart recorder. The digitized values are stored in the computer memory and also recorded on magnetic tape for later processing on the CDC 6600 computer at the National Center for Atmospheric Research (NCAR).

The computer is a PDP 8/L with extended memory. It has a memory of 8192 twelve-bit words and a 1.6-μsec cycle time. It can be used with a wide variety of peripheral equipment. In addition to the magnetic tape recorder, we use a teletype—chiefly for loading programs—and an X-Y recorder for plotting the interferograms and spectra. A crystal clock provides signals for timekeeping; for locking the chopper, which is run by a stepping motor; and for timing the interferometer advance. The control panel displays the status of the computer and allows the operator to select program options and operating modes, or even to modify the computer program while the computer is running. Digital displays show the time, the inter-

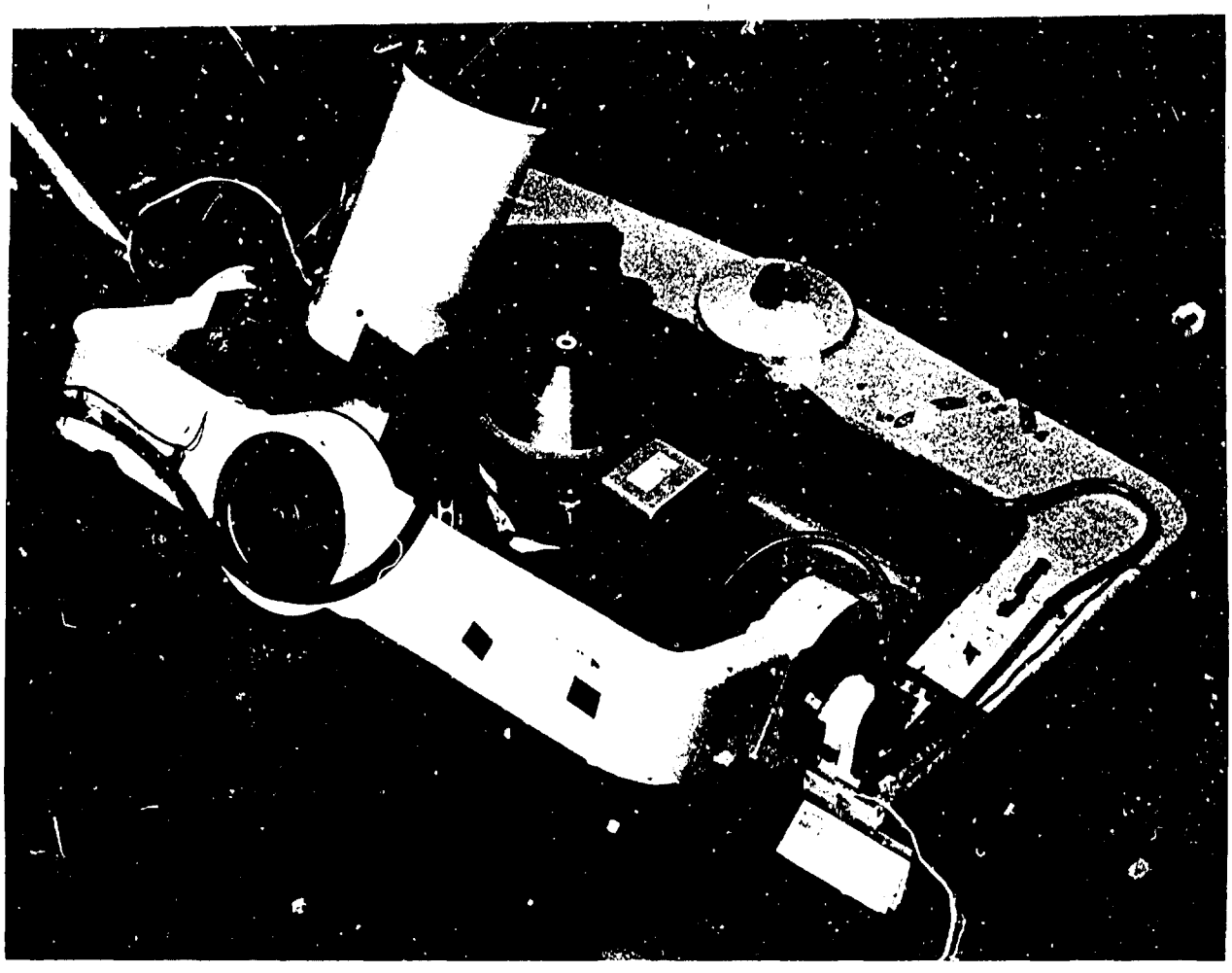


Figure 26-3. Instrument Mounted on Its Gyro-controlled [illegible] Gimbal Support

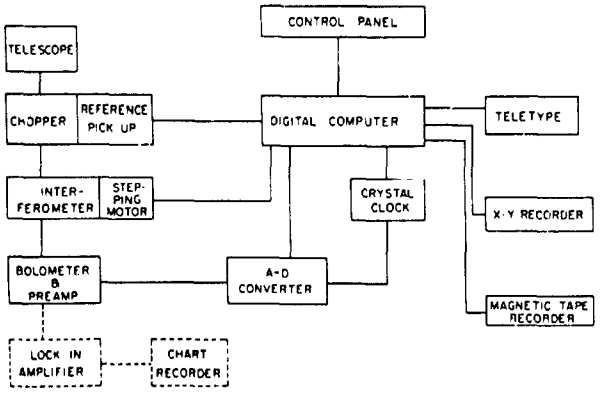


Figure 26-4. Block Diagram of the Data-handling System Used With the Interferometer

ferometer position in its scan, and an identifying sequence number.

In addition to the ability to perform programmed computations, the computer has two other modes that make it useful for controlling the experiment: the data-break facility, and the program-interrupt facility.

The data break is a method whereby numbers can be stored in specified locations of memory or retrieved from memory for output without interrupting the flow of computation in the computer. The data goes to or from memory directly, without affecting the active registers, the parts of the computer that actually perform the computation. Thus, for example, the computer can be collecting new data in one block of memory while processing data previously collected and stored in a different block of memory. The data-collection operation does not interfere with the data-processing operation.

The second facility is the program interrupt. This is a method whereby the program being executed by the computer can be interrupted by some external

device. The computer jumps to a different program to perform whatever service the external device requires; when it has completed that service, it returns to the program it was previously executing and picks up where it left off. This allows the computer to use its time efficiently, doing useful computations on one program while waiting for orders to execute another program. Input and output devices are typically much slower than the computer; the program interrupt allows the computer to switch to another job while waiting for an I/O device to catch up. It also allows the timing of the computer operations to be locked to external devices.

A description of the measurement cycle at one interferometer position should make the operation of these two special functions clear. The chopper reference pickup supplies a pulse to indicate that radiation from one beam is incident on the interferometer; the A–D converter begins digitizing the detector output at a 1-kHz rate. The digitized signals are stored in a preset buffer region of computer memory via the data-break facility without disturbing the computer operation. The computer might be doing something quite different, such as plotting the previous point on the X–Y recorder. When a specified number of samples have been recorded, the A–D converter interrupts the computer. The computer then turns the A–D converter off and adds up the digitized signals to obtain a value for the signal from that beam. When the chopper has moved to the position where the other beam is incident on the interferometer, A–D conversions begin again and the process is repeated. When 8 complete cycles of this have been performed, the computer adds all the values from one beam, subtracts the values from the other beam, thereby performing a digital synchronous rectification, plots the point, and stores it as one point of the interferogram. The computer then determines whether the interferometer has reached the limit of its travel. If it has not, the computer advances it one step through a programmed acceleration and deceleration, and the measurement cycle is completed. If the interferometer has reached a limit, the computer takes the Fourier transform of the interferogram and plots the resultant spectrum. Four seconds is required to take both the sine and cosine transform of a 512-point interferogram using the Cooley-Tukey algorithm. An additional 10 to 25 sec is required to plot the spectrum, depending upon how much structure there is in the plot. Within 30 sec after completing the interferogram, we have a complete spectrum plotted.

The interferometer takes 55 msec to advance and settle to its new position; 111 msec is then required to perform the measurement cycle, resulting in a bandpass of 4.2 Hz centered at the chopping frequency of 72 Hz. This allows measurements at the rate of 6 interferometer positions per second. We require 90 sec to obtain a 512-point double-sided interferogram. This speed introduces a problem that we call the 'dc interferogram.' Because the interferogram of the emission of sky, mirror surfaces, etc., on one side of the interferometer differs from the complementary interferogram of the emission of the cold detector on the other side of the interferometer, an interferogram of large amplitude is generated by moving the interferometer. This interferogram is not due to the difference in radiance of the two input beams but due rather to the average radiance of the two beams. It can be detected by a dc detector. Figure 26-5 is an example of this dc interferogram; it shows the output of the detector while the interferometer is scanned continuously. The small variations at the chopping frequency represent the desired signal, due to radiance differences of the two beams, and the large amplitude curve is the undesired dc interferogram.

Ordinarily, after advancing, since the interferometer sits at constant path difference for a time that is long compared to the time constant of the detector, the spike caused by the change in the dc interferogram dies out and only the chopped signal persists. But at the high stepping rate and low signal radiance we use, the signal due to the derivative of the dc interferogram is too large to drop out.

The dc interferogram consists of only a few large fringes around zero path difference since it is due to a smooth spectrum. To eliminate the signal from this source, we programmed the computer to wait an extra 111 msec before accepting data after advancing the interferometer if the interferometer was within 20 steps of zero path difference. This extra time allowed the large spike to die out. The fact that this required only a simple change in the computer program rather than a change in the hardware illustrates the power of the computer as an experiment control device.

The other serious problem that we encountered was in balancing the thermal emission of the instrument in the two beams. Since we were looking for a very low radiance source, small differences in the emissivity of the optics in the two beams were quite serious. An emissivity difference of 0.01, for example, produces a spurious signal approximately 10 times the predicted coronal signal. By careful cleaning of the

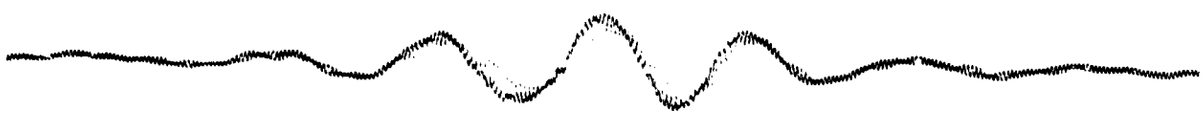


Figure 26-5. Output of Detector With Interferometer Scanned Continuously, Showing Chopped and DC Interferograms

optics in the two beams and by adjustment of the alignment, we were able to balance the total power in the two beams to within the sky noise. A spectral imbalance remained, however, due to differences in the spectrum of the spurious emission in the two beams. This spectral imbalance set the limit on our spectral sensitivity to faint sources.

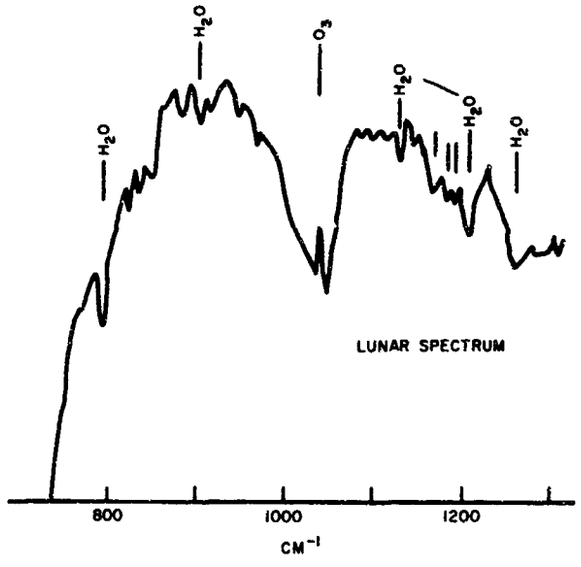


Figure 26-6. Spectrum of Region of Moon Near the Crater Kepler

Figure 26-6 shows a spectrum of the moon taken with this system. One beam was pointed to an area near the crater Kepler, the other to the sky 1.5° away. The age of the moon was 22.2 days, so the average temperature over the field of view was about 335° K.[1]

The overall shape of the spectrum is due to the transmission of the system, primarily the interference filter in the dewar. The ozone band at 9.6 μm and several minor water bands are clearly visible. The quality factor q defined by Connes[2] is approximately 3×10^5 for this spectrum; the limit is set by sky noise.

Figure 26-7 shows a measurement (at a fixed optical path difference) of sky noise—that is, fluctuations in the thermal emission of the sky—probably due primarily to fluctuations in the quantity of water vapor in the beam. This sets the ultimate limit on our radiometric sensitivity, at approximately 5 times the detector noise. At the time of the observation the sky was thermally active, with orographic clouds common. The altitude was 2700 m. The large fluctuations near the righthand side of the figure are due to wisps of clouds that are drifting across the field of view. The rms short-term fluctuations in the sky radiance are about 4×10^{-7} W/cm^2-sr-μm averaged over the 8- to 13-μm bandpass of the system. The slower fluctuations do not cause noise in the spectral region observed.

If we assume a zenith transmission of 90 percent over the spectral region and a mean sky temperature of 260° K, then the fluctuation represents approximately 1 percent of the sky radiance.

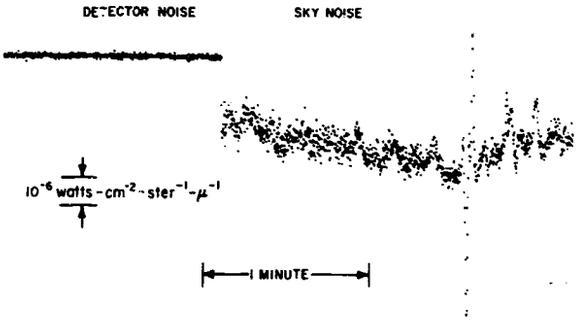


Figure 26-7. Comparison of Noise From Detector and Fluctuations in Thermal Emission of Sky

Acknowledgments

The programming of the computer was done by William D. King, now of McDonald Observatory.

This work was supported by the National Science Foundation.

References

1. Shorthill, R. W., and Saari, J. M. (1965) *Ann. N.Y. Acad. Sci.* **123**: 722.

2. Connes, J., and Connes, P. (1966) *JOSA* **56**: 896.

Discussion

Q. (K. H. Olsen, Los Alamos Scientific Laboratory): What was the field of view on the corona?

A. (W. G. Mankin): Two fields of view, each about 5 arc min in diameter, separated by 1.5°.

Q. (R. A. Hanel, Goddard Space Flight Center): Have you obtained spectra of the solar corona?

A. (W.G.M.): I am unable to say until we analyze the data. We did obtain data, but cannot yet say how much of the signal is coronal.

Q. (P. L. Richards, University of California): Were your computations done in single or double precision?

A. (W.G.M.): Computations on the computer in the system were single precision (12 bit), but data were recorded on magnetic tape in double precision for later processing on a CDC 6600 computer.

Contents

27. Fourier Spectroscopy Applied to Field Measurements

G. W. Ashley and A. G. Tescher
General Dynamics
Pomona, California

Abstract

Fourier spectroscopy has been used in a number of field measurement programs conducted by the Pomona Division of General Dynamics. The support equipment and techniques developed to achieve a measurement capability in this area are discussed. In particular, the following areas are covered in some detail: (1) field support equipment, (2) measurement instrumentation, (3) data acquisition procedures, (4) data recording system, and (5) data reduction techniques. Examples of specific data and the presentation formats used are included.

27–1 INTRODUCTION

The Physics and Infrared Section of General Dynamics, Pomona Division, has been engaged in various measurement programs utilizing Fourier spectroscopy. The program objectives have included tactical military targets as well as study of the optical properties of backgrounds and the atmosphere. High spectral resolution data requirements coupled with limited target observation time has dictated the use of multiplex spectrometers. In addition, since the majority of the targets cannot be measured in a laboratory environment, high emphasis has been placed on achieving field mobility and compact, rugged instrumentation.

27–2 FIELD INSTRUMENTATION

A complete measurement facility has been housed in a Cortez Van. Besides the basic measurement instrumentation, the van contains a 10-kW power generator, a search and range radar, VHF/UHF communications, an IRIG time code generator and control unit, a hybrid digital/analog recording system, a real-time 500-point spectrum analyzer for data assessment, and associated monitor and calibration instrumentation. The measurement instrumentation, some of which is shown in Figure 27-1, includes IR

Figure 27-1. Infrared Thermal Imaging and Interferometer Instruments

interferometers, IR thermal imaging instruments, a UV prism spectrometer, and various IR and UV radiometers.

Figure 27-2. Interferometer and Boresighted X-band Range Radar

Figure 27-2 shows one of the interferometers fitted with 12-in.-diam collecting optics. The instrument is manually pointed at the target source. Target range is automatically determined from the X-band radar, which is boresighted to the spectrometer.

27–3 INTERFEROMETERS

Figure 27-3 is a photograph of one of the Michelson interferometers and its associated electronics. The instrument was built at General Dynamics and includes a servo-controlled mirror drive system developed by Idealab. All of the instruments presently used at General Dynamics are of the rapid-scan variety, producing up to 20 scans per sec. For maximum time efficiency, a triangular mirror sweep motion

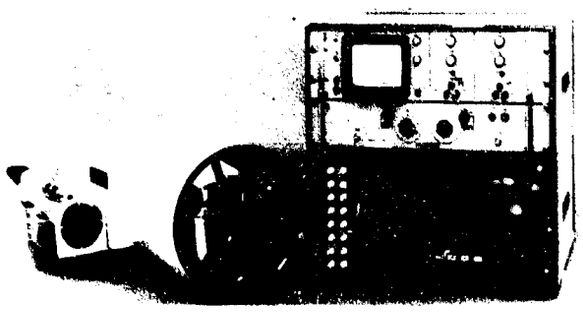


Figure 27-3. Interferometer With Control and Monitor Electronics

is used, producing identical interferograms in the two scan directions. The actual mirror sweep length of this instrument is adjustable from 0.5 mm to 1.0 cm. The instruments are fitted with interchangeable foreoptics which range from 4 to 12 in. in diam. The choice of external optics along with a wide choice of detector sizes provides instrument fields of view ranging from 1.5 mrad to 10°.

A single optical cube is used for the information channel, a laser reference, and a white light sync pulse as shown in Figure 27-4. The sync pulse is used

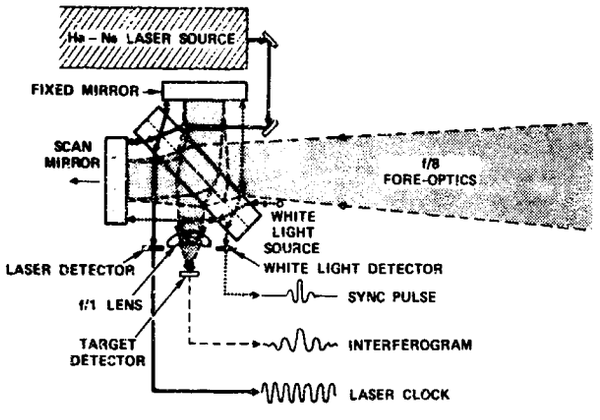


Figure 27-4. Optical Layout of Michelson Interferometer

to locate the exact center of the interferogram for coherent averaging of successive data. The laser provides the digitizing clock and wavelength reference. Combining all three functions into a single cube simplifies optical alignment of the instruments and provides a maximum of compactness.

The instruments are generally used with f/8 foreoptics and a 2-mm sweep length for the lower resolution work, corresponding to about 5 cm^{-1}. An f/1 detector lens and an f/1 cold-shielded detector operating at liquid N_2 temperatures form the remainder of the system.

A summary of the specific characteristics of two of the most frequently used interferometers is presented in Table 27-1.

One of the more difficult problems associated with the use of interferometers in a field environment is the ability to maintain optical alignment over an extended period of time. Alignment problems can be alleviated somewhat through the use of angle-insensitive corner reflectors or similar devices; however, this usually increases the weight and size of the instrument by a considerable amount. Figure 27-5 shows part of a new Michelson interferometer being developed at General Dynamics. It features a piezoelectric mirror alignment scheme coupled with three laser reference beams and phase-sensitive comparators. Included in the photograph is one of the

Table 27-1. Specific Characteristics of Two General Dynamics Interferometers

	Unit No. 1	Unit No. 2
Spectral Coverage	1.3 to 5.5 μm	3 to 16 μm
Scan Rate	20/sec maximum	20/sec maximum
Spectral Resolution	0.5 cm^{-1} to 20 cm^{-1}	0.5 cm^{-1} to 10 cm^{-1}
Field of View	1.5 mrad to 10°	3 mrad to 10°
Detectors	InSb (77° K) InAs (77° K)	Hg–Cd–Te (77° K), Thermistor (300° K)
Beam Splitter	Si on CaF_2	Ge on NaCl
Maximum Sensitivity	1×10^{-13} W cm^{-2}-μm	5×10^{-13} W cm^{-2}-μm

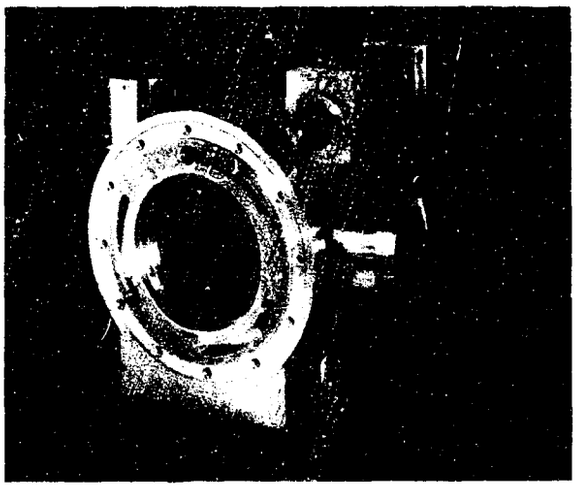


Figure 27-5. Optical Cube of Interferometer Using an Optical Alignment Servo System

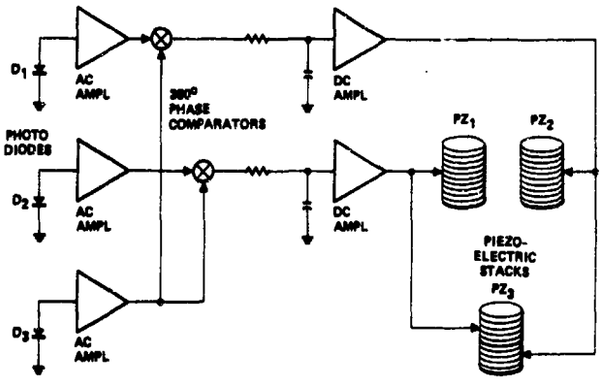


Figure 27-6. Alignment Servo Block Diagram for an Interferometer Cube

piezoelectric transducers used to position the non-driven mirror.

The laser is split into three beams and arranged to form an equilateral triangle centered about the main information channel. The output of each beam, which is modulated by the scanning interferometer, is detected by a silicon photodiode. As shown in Figure 27-6, a phase comparison is made between two of the channels, with the third used as a reference. The phase comparators provide an output voltage that is proportional to the phase difference over a ±360° angular error. The resulting dc voltage developed by any phase error, which corresponds to a misalignment of the optical cube, is amplified and applied to the appropriate piezoelectric transducer. The transducer is capable of correcting angular errors up to ±150 arc sec.

The system maintains optical alignment over extended periods of time and also during the scan, eliminating optical apodization of the interferogram. In addition, the mirror alignment system relaxes the rigid requirements normally placed on the mirror drive system. A piezoelectric mirror drive system is also being developed for this particular interferometer in an effort to completely eliminate all effects introduced by the vibration encountered in field environments.

Another interesting feature of this cube is its ability to produce modulation of the path difference. This is accomplished by superposing a small high frequency on the dc voltage that controls the piezoelectric transducers. With this technique, the interferometer can be scanned or stepped at a very slow rate without being subjected to the problems introduced by atmospheric turbulence.

27–4 DATA ACQUISITION PROCEDURE

The data acquisition and reduction procedure is summarized in the block diagram in Figure 27-7. For each series of target or background data, a calibration blackbody is also recorded. The analog interferogram is converted to digital form and recorded on a wide-band FM tape recorder. Additional data such as IRIG timing, radar range, and meteorological information are included on other channels of the recorder. At the conclusion of the field measurements, the analog tape is decoded and played into an IBM 1800

computer. Portions of the data to be analyzed are automatically selected by the IRIG search and control unit. The 1800 computer produces a digital tape that is compatible with the larger CDC 3400 used for the

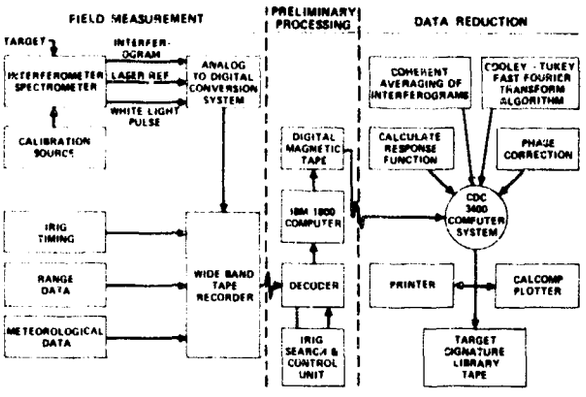


Figure 27-7. Data Acquisition and Reduction Procedures Block Diagram

actual data reduction process. Compensation for the nonlinear response of the optical and electrical equipment is achieved by comparing the uncompensated spectrum of the calibration blackbody with Planck's equation for blackbody emission. The instrument response curve derived by this process is used to compensate all subsequent spectra.

The correspondence between frequency and wavelength is established through the laser reference and used to correctly plot the abscissa scale for all data during the run. The ordinate scale is automatically determined for the calibration and the data by providing the computer with information relating to the calibration source. The calibration blackbody and instrument response curve can be plotted with the data and used to analyze the performance of the instrument. The remaining data on the tape are Fourier-transformed, corrected for the instrument response, assigned abscissa and ordinate scales, and recorded on a plotting tape. The plotting tape is used in a digital plotter to produce the final spectral signature. The final spectrum is also stored on a library tape, which is available for subsequent applications.

27-5 HYBRID DIGITIZER SYSTEM

Wideband Fourier spectroscopy places very strong requirements on the dynamic range of the recording system. Even for low-resolution spectroscopy, the dominant source of noise is tape noise if analog recording is used. To eliminate this source of noise, the raw output of the instrument should be immediately converted to digital form and then recorded.

The logical decision would indicate the requirement for a pure digital system whose output would be a computer-compatible tape. Most field data, however, is generated at rates up to 10 kHz, and the instrument is capable of even higher rates. The digital system would therefore have to operate at 30 kHz or higher. The building of such a system for field use is presently not feasible because of high cost and the large space requirements.

The hybrid system that has been adopted does avoid previous disadvantages at the expense of introducing an additional processing step in order to obtain a computer-compatible digital tape. The flow chart of data acquisition is shown in Figure 27-8.

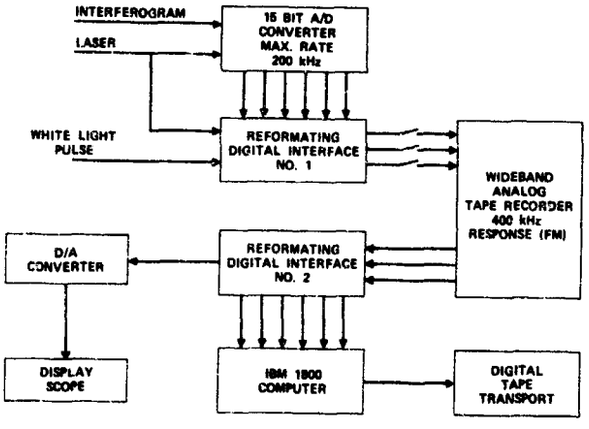


Figure 27-8. Hybrid Digital/Analog Recording System Block Diagram

The center three boxes, the A/D converter, and the two digital interfaces, are packaged together as shown in Figure 27-9. The input to the system consists of the interferogram, the laser-produced sine wave that serves as the external clock to the A/D converter, and the white light pulse. The interferogram is sampled by a 15-bit A/D converter whose output is transferred to interface unit No. 1. The white light pulse is directly encoded into a voltage level detector in the first interface. The binary output of the level detector and the 15-bit sample word constitute the 16 bits of binary information. Interface No. 1 converts the binary data into pulse-code-modulated (PCM) form and this information is recorded on a Hewlett-Packard wideband tape recorder. Out of a possible 14 channels, 3 are used for FM. The PCM data (see Figure 27-10) are immediately reread and passed through the second interface, reconverted into an analog signal, and made available for monitoring purposes. Consequently, a real-time assessment can be made on the quality of the recorded data.

Figure 27-9. Hybrid Digitizer and Associated Electronics

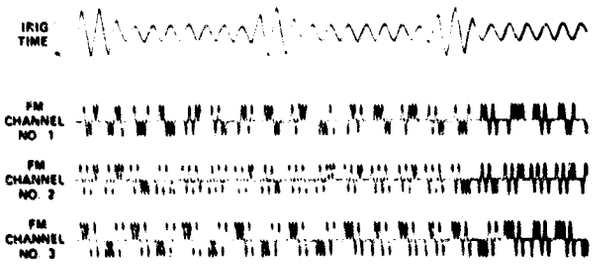


Figure 27-10. The Three-Channel PCM Format for the Digitized Data

27-6 DATA PREPROCESSING

Considerable care must be exercised to ensure that interferograms having errors are removed before coherent averaging or subtraction is performed. In particular, two types of errors should be compensated: (1) shifting of the whole interferogram due to the loss of a sample point, or (2) a digital error developed by incorrect digital recording. Unfortunately, neither error is easily observable in the interferograms. To facilitate the removal of these errors, a correlation algorithm is used. Normally, the first interferogram of a set to be averaged is designated as a reference. The first and last 500 data points of the reference interferogram and the equivalent portions of all other interferograms are used to generate a 5-point symmetric cross-correlation function.

A typical example for six interferograms is shown in Table 27-2. The units are in terms of the A/D converter. The first line is the autocorrelation function and the center value is the variance; consequently, the rms value of the interferogram at this region is approximately 30 out of a maximum possible value of 16,583. Assuming additive white noise, the rms noise level can be obtained from the square root of the difference of the variance and the peak value of the cross-correlation function, which is approximately 6. In this particular example, a 1-point shift is observed in some interferograms.

In a second computer program, the interferograms are shifted to bring them into coincidence and then sequentially subtracted from their reference. Any difference between interferograms that exceeds a preset level based on the rms noise level, is printed out. If a digital error is discovered, it is either replaced by the average of two adjacent points or the entire interferogram is excluded from the averaged set, depending on the particular test situation.

27-7 SELECTED EXAMPLES

Figure 27-11 shows 4- to 4.8-μm spectra of a calibration blackbody source taken at 5 ft and at 100 ft, respectively. The observation time was 5 sec. The resolution is better than 1 cm^{-1} and the observable noise level is negligible.

The spectrum of the hot exhaust of a typical jet engine is shown in Figure 27-12. The aircraft was located at 200 ft from the observation point. The atmospheric absorption lines are superposed on the continuum radiation of the hot CO_2 gas. Again, the spectral resolution is better than 1 cm^{-1}.

The last example, Figure 27-13, shows the spectral distribution of two spatially related sky backgrounds and their differences. The spectral contrast or difference plots, which are plotted in both linear and logarithmic form, are obtained by subtraction at the interferogram level. Data of this type are used for system optimization in the development of infrared seeker and search devices. The resolution of this data is on the order of 5 cm^{-1}.

27-8 CONCLUSIONS

Fourier spectroscopy can be successfully used in field environments, but special precautions must be taken to ensure data of high quality. Solutions to some of the field-generated problems have been presented in this paper; however, more effort must be made in the area of instrument development. In

Table 27-2. Five-Point Correlation of Six Interferograms

Interferogram No.	Correlation Position				
	−2	−1	0	+1	+2
1*	19	691	994	694	22
2	−10	654	934	669	14
3	686	958	677	5	−661
4	657	931	648	−15	−667
5	687	962	680	7	−651
6	46	677	910	596	−76

Correlation at beginning of interferogram

Interferogram No.	Correlation Position				
	−2	−1	0	+1	+2
1*	29	700	1002	701	29
2	−19	676	982	721	50
3	700	993	715	27	−667
4	659	951	682	9	−677
5	699	951	644	−39	−698
6	81	736	972	642	−55

Correlation at end of interferogram

* Interferogram No. 1, used as reference

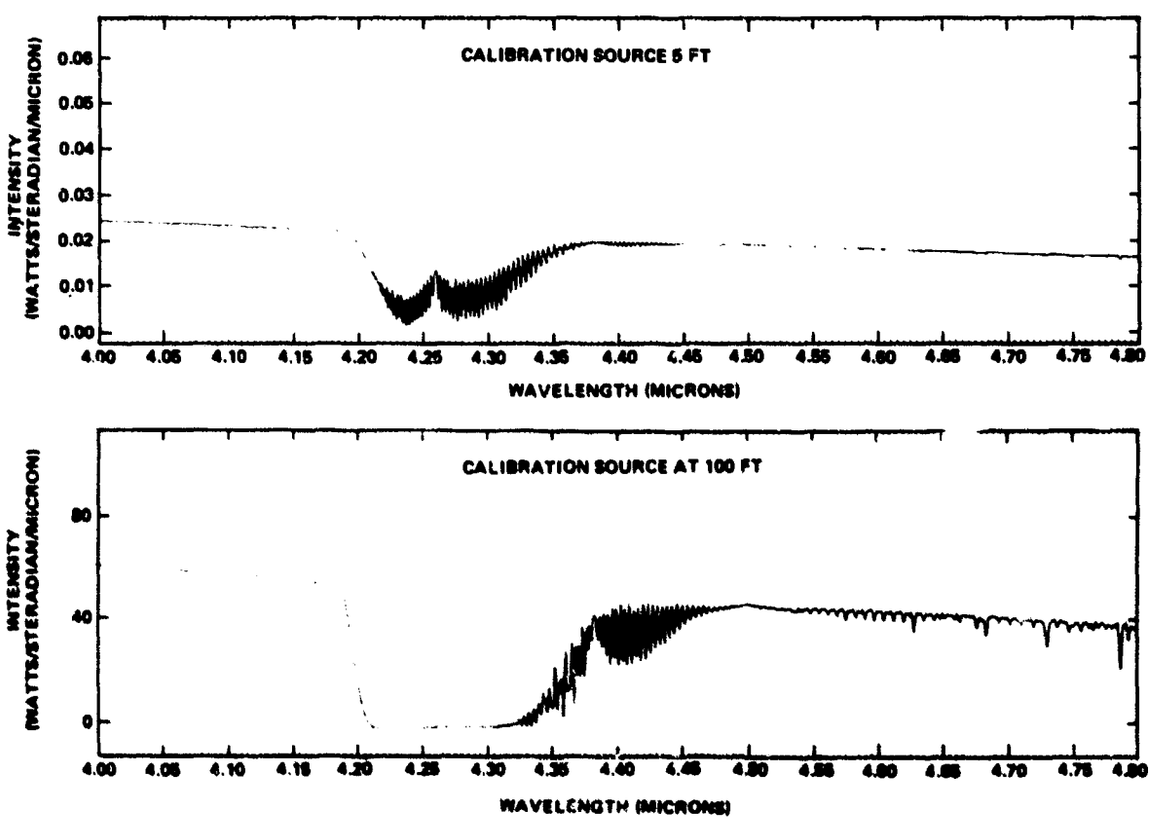


Figure 27-11. Spectrum of a 1000° Blackbody Taken at 5 ft and at 100 ft

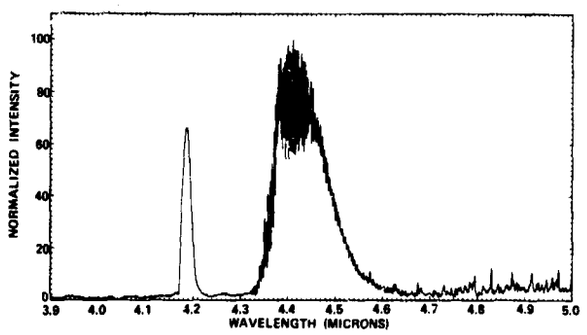


Figure 27-12. Spectrum of the Hot Exhaust of a Typical Jet Engine

particular, field-encountered vibrations, especially those introduced by airborne measurement platforms, can be detrimental to the satisfactory performance of interferometers as they are now designed. Another area requiring attention is the development of a true real-time high-resolution data reduction facility that can be used directly in the field. As new, highly sophisticated Fourier systems are developed, particular attention must be devoted to operational reliability.

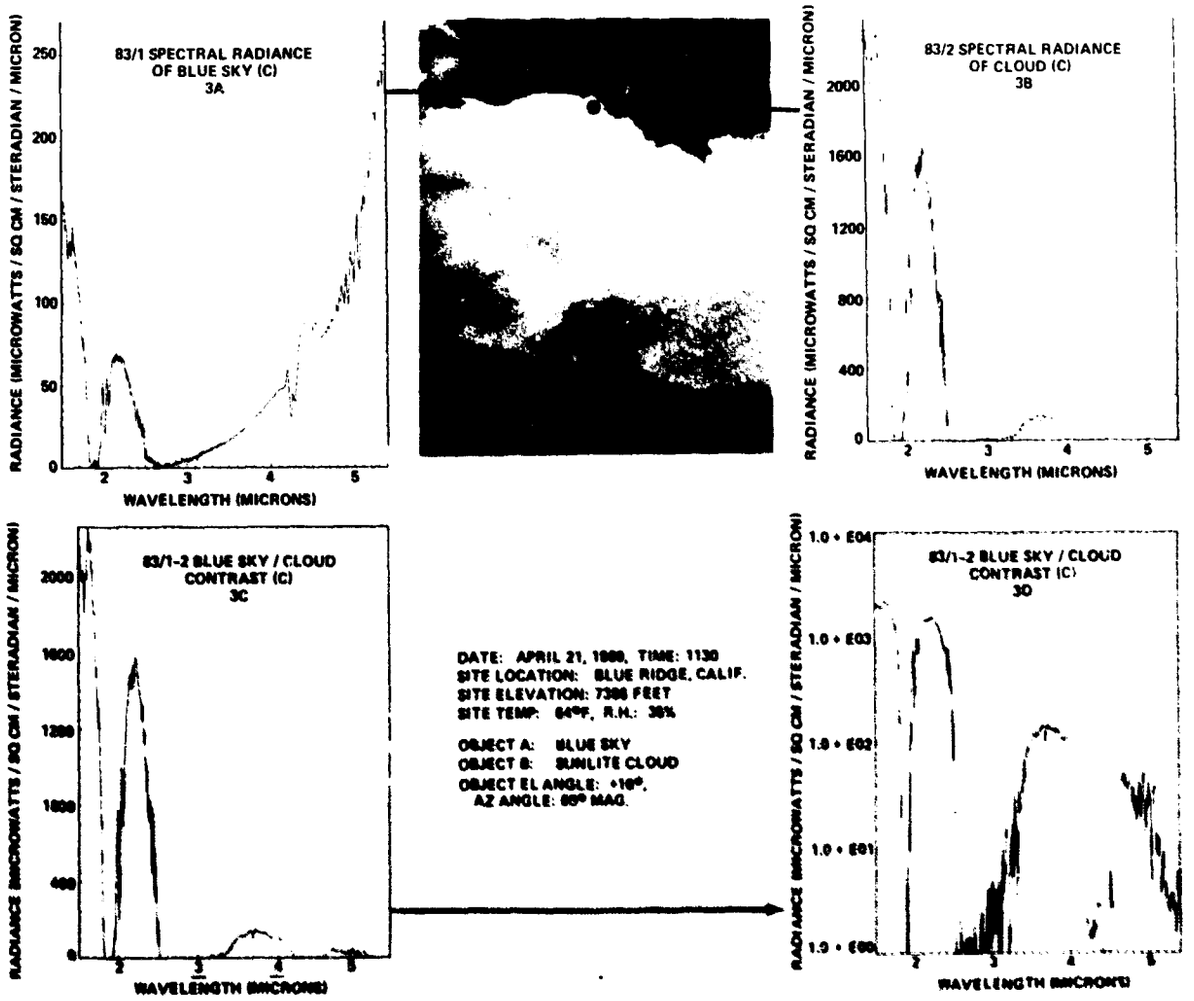


Figure 27-13. Spectral Distribution of Two Spatially Related Backgrounds and Their Differences

Discussion

Q. (John Rex, AFCRL): How do you prevent your 3-phase laser parallelism circuit from inadvertently locking-in on the wrong fringe?

A. (G. W. Ashley, General Dynamics): Each photodiode is actually a bicell consisting of two 0.075-in. by 0.150-in. detectors separated by a 0.005-in. gap. Each of the three 2-mm laser beams is split into two equal parts at the detector. The phase between the halves of the beam is measured in a manner similar to that discussed in the text, and is used to reduce any large misalignment to a point where all three beams are within a 360° phase error.

Contents

28. A Liquid-Helium-Cooled Michelson Interferometer

Gordon C. Augustson
Ames Research Center, NASA, Moffett Field, California 94035

Niels O. Young*
Block Engineering, Inc., Cambridge, Massachusetts

Abstract

Many objects of great interest to astronomers have emission spectra in the far infrared spectral region. In order to study these objects, an instrument with high throughput, high sensitivity, and high spectral resolution is required. To achieve maximum sensitivity, an infrared detector must be cooled to liquid helium temperature and be rugged enough to operate after a rocket launch places it above the earth's atmosphere. Because of its compactness and mechanical configuration, a Michelson interferometer is the easiest dispersing device to cool, although severe problems are encountered. The solution of these problems is discussed in this paper, which describes the construction of a laboratory prototype liquid-helium-cooled Michelson interferometer with 1 cm^{-1} resolution, to be launched on an Aerobee rocket.

28–1 INTRODUCTION

Many problems in astronomy remain unsolved because they require knowledge of the spectral characteristics of various astronomical objects in the far infrared spectral region. These problems are concerned with the composition of the interstellar medium, its cooling and heating mechanisms, and its dynamic processes. These subjects relate, in turn, to the formation of protostars, the nature of the galactic nuclei, and ultimately to extragalactic objects such as Seyfert galaxies and quasars. To study these problem areas requires instrumentation capable of detecting and spectrally characterizing radiation from weak, discrete emission sources in the spectral region 10 to 1000 microns. This paper describes the design and construction of such an instrument: a liquid-helium-cooled Michelson interferometer which is intended ultimately to be flown on an Aerobee 150 rocket.

* Now Consulting Physicist, Old Winter Street, Lincoln, Mass. 01773

28–1.1 Astronomical Objectives

The emission spectra of interest to astronomers are expected to come from two main types of sources: molecular hydrogen and the neutrals and ions of C, N, O, Ne, Mg, Fe, Si, and S. The emission from molecular hydrogen will be discussed first.

Determining the abundance of molecular hydrogen is of great importance to astronomers because of the effects that appreciable amounts of hydrogen have on star formation, the temperature balance of the interstellar medium, and the total mass content of the galaxy. Because of its dipole nature, the only allowed transitions for molecular hydrogen are electronic transitions, which occur in the ultraviolet (Spitzer, Dresler and Upton, 1964).[1] It should, however, be possible to detect the quadrupole rotational transition $pH_2(J=2) \rightarrow pH_2(J=0)$ at $\lambda = 28.17\ \mu$. This radiation is expected from interstellar clouds in the galactic plane (Takayanagi and Nishimura, 1961)[2] and from collapsing protostars in the galactic plane (Gould,

Table 28-1

Radiating Species	Location of Source	Wavelength (μ)	Radiance (W/cm^2-sr)	Flux at Earth (W/cm^2)
H_2 molecule	Interstellar clouds in the galactic plane*	28	4.7×10^{-11}	
H_2 molecule	Collapsing protostars in the galactic plane†	28		3×10^{-16}
CII	Shockfront between interstellar clouds**	156	10^{-13}	
OI	Shockfront between interstellar clouds**	63	10^{-14}	
OIV	Planetary nebula†† NGC 7027	26		4×10^{-15}

* Takayanagi and Nishimura (1961)[2] with hydrogen density $n(H_2) = 10^{-2}$ cm^{-3}
† Gould (1965)[3]
** Field, Rather, Aanstead, and Orzag (1968)[4]
†† Delmer and Gould (1967)[5]

1964).[3] The expected radiance of these sources and their flux at the earth are presented in Table 28-1.

Emission from the forbidden fine-structure magnetic dipole transitions of the ground states of the neutrals of C, N, O, Ne, Mg, Fe, Si, and S occurs in the far infrared spectral region. This radiation is considered to be an important source of cooling of interstellar clouds that have been heated by inelastic cloud-cloud collisions (Field, Rather, Aanstead and Orzag, 1968).[4] Radiation from these species is also an important source of infrared emission from planetary nebulae (Delmer and Gould, 1967).[5] The flux expected at the earth from these sources is listed in Table 28-1. Gillett and Stein (1969)[6] found that the NeII line at 12.8 μ of the planetary nebula IC 418 has a flux of 2.1×10^{-17} W/cm^2 at the earth with a continuum flux of 6.1×10^{-17} W/cm^2-μ. This last observation has demonstrated the value of infrared spectral measurements in determining the composition and physical conditions in planetary nebulae. Petrosian and Bahcall (1969)[7] have examined and rejected the possibility that redshifted emission from these species might be responsible for the universal 3°K background radiation. This type of emission may, however, be responsible for the large IR fluxes observed in the center of our own galaxy.

28-2 CHOICE OF INSTRUMENTATION

From Section 28-1.1 it can be seen that the objects of interest are faint sources. Even at balloon altitudes, atmospheric radiation is many orders of magnitude greater than radiation from the astronomical objects. Also, the opacity of the earth's atmosphere renders the greater part of the far infrared spectral region inaccessible to observers from the ground.

Because of the faintness of the sources, detectors of the highest sensitivity are required. Theoretically, the sensitivity of a detector exposed to background radiation at its own temperature is proportional to $T^{\frac{5}{2}}$ (Putley, 1964).[8] It is therefore extremely important to maintain such a detector at the lowest possible temperature and to exclude all possible sources of external radiation. Thus, to prevent spectral and thermal contamination from the earth's atmosphere it is important to make the desired observations above the atmosphere. Although the total observing time is relatively short (approximately 5 min), a sounding rocket is an acceptable vehicle for many of the observations discussed.

High spectral resolution is required to detect and discriminate the emission lines against the background radiation. Since most of the objects discussed are interstellar clouds whose random velocities of about ±6 km/sec would cause Doppler broadening of the emission lines, no lines of widths less than approximately 0.004 cm^{-1} should be expected. In the case of radiation from molecular hydrogen, Takayanagi and Nishimura (1961)[2] have shown that for a density $n(H_2)\geq10^{-2}$ cm^{-3}, a resolution of about $\Delta\nu=1$ cm^{-1} is sufficient to detect the emission lines in the presence of the expected thermal bremsstrahlung background.

High throughput θ ($\theta=\Omega S$, where S is the area of the aperture or entrance pupil, and Ω is the solid acceptance angle) is required because of the extent of the sources and to compensate for the poor pointing accuracy of the rockets. Grating instruments lack throughput, require a large volume, and have severe order-sorting requirements in the spectral region of interest. Both double-beam interferometers (typified by the Michelson interferometer) and multiple-beam interferometers (typified by Fabry-Perot interferometers) are compact and have a large throughput (Jacquinot and Girard, 1967).[9] For a given throughput, the half-angle of rays is the same whether crossing the Fabry-Perot or the Michelson. The entrance diameter will therefore be the same for both. Because of their similarities, both types of instruments were considered for the observations.

The greatest advantage of a Michelson is that it can be used for spectral scanning. It can also, as will be discussed later, be used for detecting discrete line emission. Its beam splitter requirements are not as severe as those of a Fabry-Perot, and it requires essentially very little predispersing. It is possible to scan a Fabry-Perot spectrometer by stacking several scanning etalons in series as in the PEPSIOS

(photoelectric etalon pressure-scanned interferometric optical system) configuration, but such an arrangement is neither simple nor compact and would be very difficult to cool (Mack, McNutt, Roesler, and Chabbal, 1963).[10]

The chief advantage of a Fabry-Perot is that it has the highest resolution of any dispersive device. The resolution of a Fabry-Perot is $\Delta\nu \approx 1/(B\mathfrak{F})$, where B is the beam retardation (twice the mirror separation) and $\mathfrak{F}$ is the finesse

$$\mathfrak{F} = \pi\sqrt{R}/(1-R).$$

The resolution of a Michelson is $\Delta\nu \approx 1/B$. Hence, if the reflectivity R of the beam splitter is low, and the finesse consequently small, the performance of a Michelson differs little from that of a Fabry-Perot interferometer. The precision of alignment and scanning motions of a Fabry-Perot is about $\mathfrak{F}$ times that required for a Michelson.

The finesse of a Fabry-Perot is determined by the reflectivity of the beam splitter, which thus becomes a very crucial component. Dichroic beam splitters have been developed for wavelengths out to 30 μ (Heinrich, Bastien, Dos Santos and Ostrelich, 1966),[11] but such beam splitters have not been made for wavelengths beyond 30 μ because no suitable materials are available.

For the longer wavelengths, Fabry-Perots featuring metal grating and grid beam splitters have been developed. Lewis and Casey (1952)[12] and Casey and Lewis (1952)[13] developed the theory for reflective metal gratings and Fabry-Perots made with these gratings. They showed that a lossless grating with $g/a = 10$, where a is the radius of the grating wires and g the grating constant, would have a reflectivity of $R = 83$ percent at $\lambda/g = 2$ and should give a Fabry-Perot of finesse $\mathfrak{F} = 16.8$. At 30 μ, a grating with these characteristics would have to have $g = 15\,\mu$ and $a = 1.5\,\mu$, which is beyond the present state of the art.

Renk and Genzel (1962)[14] have made nickel grids with $g = 50\,\mu$ and $a = 6\,\mu$. These grids have a reflectivity of $R = 56$ percent at $\lambda/g = 2$ and should give $\mathfrak{F} = 5.3$, but interference filters made of these grids do not achieve the expected finesse because of mechanical support problems.

Renk and Genzel (1962)[14] derived approximate formulas for the finesse of a Fabry-Perot having metal grids and showed that $\mathfrak{F} \propto \lambda^2$. Although the finesse would hence be expected to increase at the longer wavelengths, it is doubtful that grids with $g = 15\,\mu$ and $a = 1.5\,\mu$ can be made at the present time. Since it is thus unlikely that a Fabry-Perot interferometer can be made to have a finesse greater than 5 over an appreciable portion of the far infrared spectral region, and suitable order-sorting filters are unavailable, the Fabry-Perot interferometer was rejected in favor of the Michelson interferometer as the most appropriate instrument.

28-3 THE MICHELSON INTERFEROMETER

The necessity of cooling the interferometer can be appreciated by comparing the radiant emittance of the instrument at various temperatures with the irradiance expected from the astronomical sources mentioned in Section 28-1.1. At 100 μ, the radiant emittance from the interferometer would be 6.4×10^{-7} W/cm^2-cm^{-1} at a temperature of 77°K, whereas at 5°K it would be 1.2×10^{-18} W/cm^2-cm^{-1}. [An emissivity of 1 is assumed in each case.] Below 5°K, therefore, radiation from the spectrometer will not be restrictive on the measurements. Also, as mentioned in Section 28-2, the performance of a thermal detector is theoretically a function of T^5; thus, if the geometry and spectral filtering are equivalent, a detector exposed to 4°K radiation would be 1.6×10^3 times as sensitive as one exposed to 77°K radiation. In practice, however, this entire gain is not realized because most cryogenically cooled bolometers are operated with cooled radiation shields and filters. Cooling the entire interferometer to 4°K places severe requirements on the interferometer alignment, beam splitter, and drive mechanism.

28-3.1 The Beam Splitter

Double-beam instruments are usually operated in the Michelson or lamellar configuration. Lamellar grating instruments are nondegenerate (and so half the radiation is not lost) but they lack throughput; they are also bulky, hard to cool, and diffraction-limited to wavenumbers less than $\nu = F/sg$ (where F is the focal length of the collimating mirror, s the slit width, and g the grating constant). To avoid these problems, the Michelson configuration was chosen.

The most critical component of a Michelson interferometer is the beam splitter. Beam splitters featuring dielectric coatings on thick substrates at wavelengths up to 30 μ have been discussed. At the longer wavelengths, Si and Ge could be used, but there would be important Fresnel reflection losses to overcome. Polyethylene 0.020-in. thick would cover the spectral range but would not remain flat over a 2- to 4-in. aperture. Solid substrates also require a compensation plate—a further complication.

Vogel and Genzel (1964)[15] have studied the transmission and reflection of metallic meshes in the far infrared as a function of polarization and angle of the incident beam. For an ideal beam splitter the product RT should be 0.25 and the efficiency would be 100 percent. For a normal unpolarized beam, a grid begins to become effective as a beam splitter at $\lambda/g > 1$. At $\lambda/g \approx 2$, a grid is effectively 100 percent efficient. At $\lambda/g > 3$ the efficiency of a lossless grating becomes less than 20 percent. When a grid is inclined at 45° and polarized light is incident, severe disruptions of the efficiency of the beam splitter occur between $1 < \lambda/g < 2$. Thus, the effective spectral

range of a grid beam splitter is reduced to approximately $2<\lambda/g<3$. The shorter wavelengths would require meshes finer than any yet fabricated. It is doubtful that such fine meshes could be stretched to maintain the desired flatness for apertures over 2 in.

Thin dielectric pellicles have many advantages. They are easy to fabricate, require no compensating plates and are almost as efficient as metal meshes (Richards, 1964).[16] Complications do arise, however, because pellicles have a finite thickness that in effect makes them solid Fabry-Perot interferometers with their own channel spectra. The transmission T of such a pellicle at a wavenumber ν is given by

$$T = M_0[\tfrac{1}{2} - \tfrac{1}{2}\cos(4\pi n x \cos\alpha\nu)], \qquad (28\text{-}1)$$

where M_0 is the modulation efficiency, n the refractive index of the material, x the thickness of the material, and α the angle of incidence. The modulation efficiency

$$M_0 = 2(R_s T_s + R_p T_p),$$

where R_s, T_s and R_p, T_p are the fraction of the intensity of the incident beam reflected and transmitted for perpendicular- and parallel-polarized light, respectively; R_s, R_p and T_s, T_p are obtained from the Fresnel formulas (Born and Wolf, 1964).[17] Where Mylar has a refractive index of 1.72, $M_0 = 0.640$. This is near the maximum possible value of M_0 (M_0 is presented here on the basis that for perfect sine wave modulation, $M_0 = 1$).

To determine if Mylar would function adequately at the low temperatures required, a program was undertaken to measure characteristics of Mylar. Pellicles were stretched by placing a sheet between two Viton O-rings, which were displaced outward as two conical facing rings were tightened. Various cements were used in attempts to fasten the stretched pellicle to a metal support ring. All were abandoned in favor of compressing the stretched pellicle between two optically flat metal surfaces. It was found that a template could be placed on the stretched pellicle to indicate the desired position of holes for clamp bolts and a light pipe, and that these holes could be melted into the mylar with a soldering pencil without affecting the figure of the pellicle or causing any tearing.

The sag under gravity for a horizontal film stretched on a 4-in.-diameter frame to about $\frac{1}{4}$ the yield strength of Mylar was found to be only about 0.3 μ, which can be tolerated even for an instrument used on the ground. The surface tension of films of different thicknesses was measured by deflating the pellicles to a given radius of curvature and then measuring the required pressure differential. Thus, $T = R(P/2)$, where R = radius of curvature, and P = pressure. Table 28-2 summarizes the findings on two films.

Table 28-2

Thickness (in.)	Pressure (psi)	Surface Tension (lb/in.)	Flatness (μ)
0.00100	3430	3.43	0.58
0.00015	6670	1.00	0.73

Each film was cooled several times to 77°K. At this temperature the surface tension increased, because the coefficient of expansion of Mylar exceeds that of the aluminum frame. Films at the low temperature were observed to be more resistant to denting by a sharp object. As soon as a film was immersed in liquid nitrogen, its surface tension decreased by an estimated 10 percent, and it remained at this level unchanged by subsequent immersions. These results were later found to be in agreement with the findings of the Air Force Materials Laboratory (1968).[18]

28-3.2 Interferometer Design Considerations

Once it has been determined that Mylar beam splitters over 4 in. in diameter could be used, work was undertaken to establish the configuration of the dewar, telescope, and interferometer. In designing the interferometer system, the throughput θ of four elements had to be matched. If S is the aperture area, then $\theta_A = 2\pi S(1-\cos\alpha_A)$, where α_A is the aperture half-angle. The resolution of a Michelson deteriorates as the incident beam is decollimated. If the half-angle of the beam is α_R, then $\alpha_R = (2/R)^{\frac{1}{2}}$, where $R = \nu/\Delta\nu$ (Vanasse and Sakai, 1967).[19] One of the limiting elements is the detector. The maximum practical size is 5×5 mm. Also, the maximum acceptance half-angle is approximately 60°. For a resolution of 100, $\alpha_R \leq 7.85°$. The pointing accuracy of a rocket spectrometer with inertial guidance is ±4°. Thus, $\alpha_A = 5°$ should be adequate. The throughput of 3.75-in.-diam. collecting optics with a half angle of 5° is $\theta_A = 1.706$ cm^2-ster. If the radiance of the molecular hydrogen cloud is 4.7×10^{-11} W/cm^2-ster and the noise equivalent power (NEP) of the detector is 10^{-14} W, then $S/N = 80.2$, allowing for a system efficiency of 0.1. Since this signal-to-noise ratio is adequate for the measurement, the interferometer may be built with the entrance aperture as the interferometer back mirror. As will be seen, this results in a very compact easy-to-cool system.

One of the severest problems in making a liquid-He-cooled interferometer is to design the scan mechanism

for the moving mirror. The mirror must be capable of moving 2 cm while maintaining parallelism of its surface. To avoid the use of bearings and lubricants, we developed a four-sided support that used spring strips and rollers to position the mirror. To test the device, a He-Ne laser was used to produce Fizeau fringes between a mirror and the beam splitter. Parallel motion was maintained over 2 cm of travel to within 5 fringes per cm. It was later determined that 'low micron size' tungsten diselenide powder made a good cryogenic lubricant, and we developed a linear bearing consisting of rows of ball bearings retained in a race surrounding a cylindrical shaft of hardened steel.

28–4 DESCRIPTION OF THE INTERFEROMETER

The dewar vessel, designed by Cryonetics Corp. and built by Janis Corp., is shown in cross section in Figure 28-1. Its dimensions were made to match

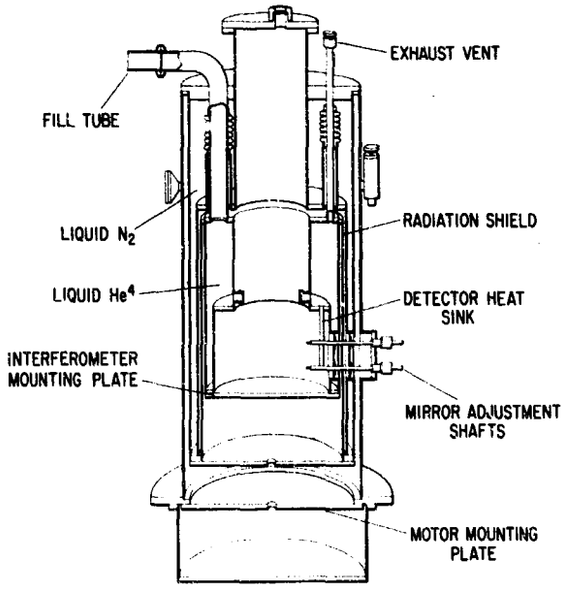


Figure 28-1. Cutaway Drawing of the Dewar Vessel (Access to the dewar is obtained by removing the bottom pan and each of the mounting plates in turn. The adjusting shafts are pulled back from the interferometer cube once adjustment has been made.)

rocket dimensions as closely as possible, with one exception: the laboratory prototype used liquid nitrogen and a vacuum space to insulate the liquid helium; in the flight model the nitrogen would be replaced with super insulation.

In the laboratory dewar, the nitrogen extends above and below the helium space. The bottom of the nitrogen space is terminated with a removable plate which shields a helium temperature plate from the motor plate. The dewar has a helium capacity of 25 l and a hold time of approximately 12 hr.

Temperature in the dewar is controlled by monitoring the overpressure and using platinum and carbon thermometers at various locations. Radiation shields inserted in the entrance cylinder are kept in thermal contact by beryllium-copper spring strips. These shields can be seen in Figure 28-2, which is a photograph of the disassembled dewar including the light baffles that are inserted in the vacuum space in front of the interferometer. One baffle is maintained at liquid-nitrogen temperature and the other at liquid-helium temperature. A light pipe passes through both baffles, with a rotating filter wheel at the cold end.

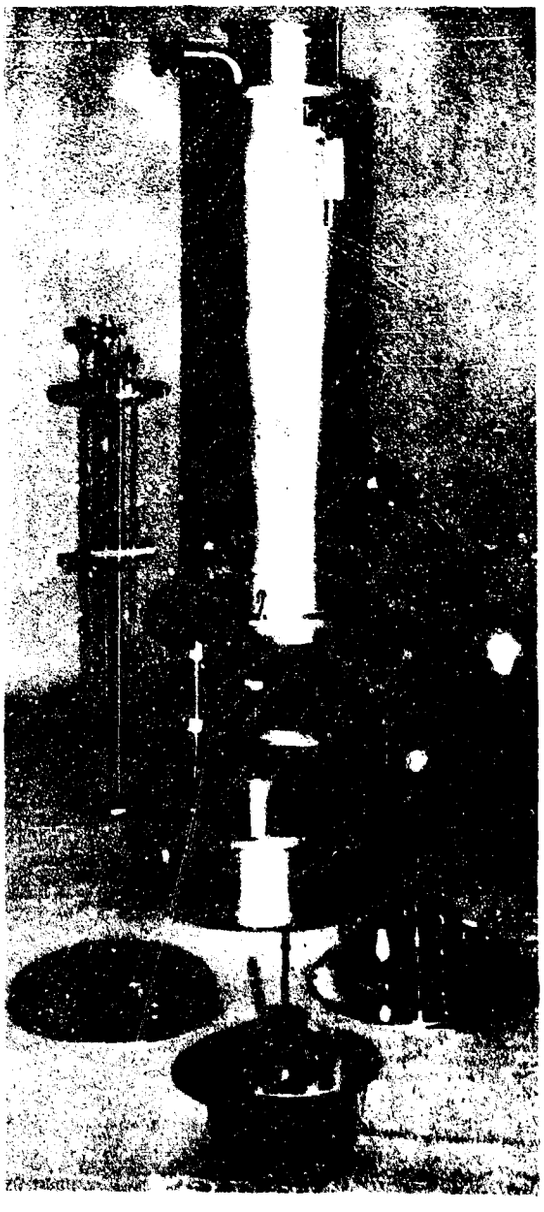


Figure 28-2. The Dewar Vessel, Disassembled

The interferometer cube is shown in cross section in Figure 28-3, and as assembled on the interferometer mounting plate in the photograph in Figure 28-4.

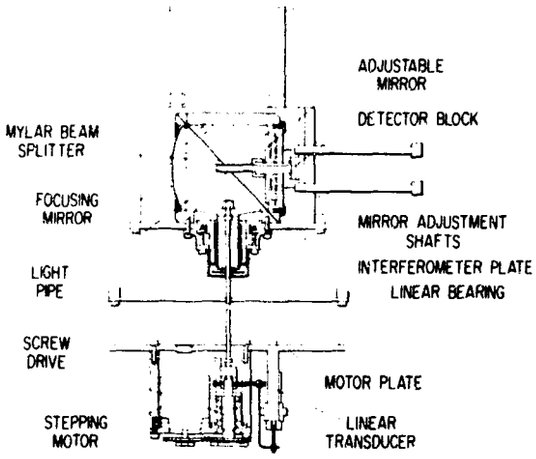


Figure 28-3. Cutaway Drawing of the Interferometer Cube (radiation enters the instrument from the top)

Figure 28-4. Photograph of the Interferometer Block (thermometer attached)

The vacuum-qualified stepping motor is geared to operate a screw that pulls on a thin stainless steel wire; the motor, its gears, and screw mechanism are attached to the room-temperature motor-mount plate. A stepping motor is used because of the ease with which it can be programed and because pulses from the stepping indexer can be used to control signal sampling. A linear-position transducer is used to indicate coarse positions of the movable mirror.

The interferometer cube is made in two almost identical sections divided at the diagonal. The interferometer, bearing support, and attaching hardware are all machined of the same type of aluminum, annealed several times by hot and cryogenic cycling before and after machining. Each half of the interferometer cube is polished optically flat across the diagonal face so that it can serve as a clamp for the Mylar beam splitter. When the interferometer cube is assembled, the pellicle is stretched and perforated as explained earlier, the two sections are bolted together surrounding the Mylar, and then they are cut from the stretcher and trimmed.

All the mirror surfaces are electroformed and gold-coated. The concave mirror that focuses the radiation is spherical and approximately $f/0.4$. The focused radiation is incident upon the surface of a cylindrical light pipe 6 mm in diameter and square in cross section at the detector end to match the 5 mm by 5 mm detector.

The light pipe, electroformed over a mandrel and then gold-plated, does not touch the detector but is butted loosely against the detector block in a way that will ensure tight contact at He temperatures. It passes through the adjustable plane mirror, which is mounted in a ball-and-socket arrangement and adjusted by means of four opposing screws placed around the periphery of the mirror mount.

The detector is a Texas Instrument Ga:Ge bolometer specially doped to perform optimally at the 3°K space background. It is mounted inside a chamber in a copper block, which is in turn attached to a copper pillar suspended from the helium dewar and mounted between four mirror adjustment shafts.

The four long shafts, provided for adjusting the mirror angle, enter the dewar vacuum space through a Viton O-ring seal. They then pass through two orifices, one cooled to liquid-nitrogen temperature, the other being in the liquid-helium heat shield. Thermal contact is made by strips of beryllium-copper spring that surrounds these orifices.

In operating the interferometer, the desired fringe pattern is achieved by adjusting the mirror-positioning screws, using sodium D light at room temperature. Next, the entire instrument is cooled to liquid-nitrogen temperature, and readjusted. The shafts are then pulled back from contact with the interferometer cube so as to eliminate a thermal leakage path. It was found that once the interferometer has been adjusted at the nitrogen temperature, it remains adjusted through the helium temperature cycle. In flight operation, the mirror would be adjusted by remote control.

The moving retardation mirror is attached to the previously described bearing and is forced toward the beam splitter by a coil spring. The screw translating mechanism is attached to the mirror by a thin stainless-steel wire that passes through the liquid-nitrogen temperature plate. The motor can be operated to accomplish two types of scan modes: a repetitive rapid

scan mode through zero retardation for scanning an entire spectrum, or a scan mode that avoids the zero retardation position when used as a single line detector.

Foskett and Weinberg (1969)[20] have shown how a Michelson interferometer can be used as a detector of discrete line radiation and how when used in this mode it can discriminate against background radiation. Such use takes advantage of the fact that the Fourier transform of an emission line is a sine or cosine function, which can be detected by means of a synchronous amplifier without the complexity of a chopper mechanism, a considerable advantage when the interferometer is cooled to helium temperature. Some complication arises because the sine wave is modulated by the apodization function, but if the instrument is not scanned through zero retardation, background radiation is strongly discriminated. The output of a Michelson is given as a function of time by

$$G_I = \int_0^\infty I(\nu) R(\nu, t)\, d\nu, \qquad (28\text{-}2)$$

where

$$R(\nu, t) = \cos\left[2\pi\nu(B_0+\phi) + 2\pi\nu x \cos\left(2\pi f_x t + \psi\right)\right], \qquad (28\text{-}3)$$

with B_0 the mean retardation, x the mean mirror excursion distance, f_x the mirror frequency, and ψ representing errors in the mirror retardation and synchronous demodulator. The Fourier expansion of Eq. (28-2) is

$$G(t) = \frac{A_0}{2} + \sum_p \left(a_p \cos 2\pi pft + b_p \sin 2\pi pft\right), \qquad (28\text{-}4)$$

where a_p and b_p are the Fourier coefficients.

If, for a monochromatic source,

$$I(\nu) = I_0\, \delta(\nu - \nu_0), \qquad (28\text{-}5)$$

then it can be shown that

$$c_p \approx 2J_p(2\pi\nu_0 x) \sin\left(p\,\frac{\pi}{2}\right), \qquad (28\text{-}6)$$

where $c_p = (a_p^2 + b_p^2)^{\frac{1}{2}}$, and J_p is the familiar Bessel function of integer order; c_p maximizes the odd harmonics that, with good phase-sensitive detection, will give the strength of the line of interest as a simple output signal.

If a complete spectral scan is desired, the instrument can be operated in a rapid-scan mode. This mode also requires no chopping. Rapid scanning yields the interferogram in a short time, thus greatly reducing the effects of pointing errors. Requirements for the amplitude capability of the telemetry can also be greatly reduced because the individual interferograms can be accumulated on the ground and then summed, and any $1/f$ noise in the detectors or electronics will be greatly reduced.

The scan rate required can be determined as follows. The fringe passage frequency f_ν at some wavenumber ν is given by

$$f_\nu = \nu(B/t)$$

where B is the retardation and t is the scan time. If $\nu = 357.1\ \text{cm}^{-1}$ at 28 μ and B is 1 cm, $f_\nu = 71$ Hz for a scan time of 5 sec. Generally, f_ν should not be greater than $2\pi/t_0$ where t_0 is the time constant of the detector (in the case of Ga:Ge bolometers, $t_0 \approx 40$ msec). Thus, f_ν should not exceed 80 Hz. A 5-sec scan time will allow 60 interferograms to be averaged during a 5-min observing period, or approximately 50 sec of observing time per spectral element, which is greater than required to attain the S/N indicated earlier.

In laboratory use the instrument was precooled with liquid nitrogen, which was removed from the helium space by pressurized dry nitrogen gas. This space was then filled with liquid helium which was pumped over by a 70 m^3/hr Heraeus pump ballasted with a 30-gal tank to prevent pump noise from affecting the detector. The pressure was reduced until the lambda point of helium was reached.

All attempts to use the Ga:Ge bolometer to obtain an interferogram at liquid-helium temperature were unsuccessful. The extreme doping given the detectors to optimize their performance with a space background caused them to saturate for all source and filter combinations tried. They functioned properly only when all external radiation sources were blocked by an aluminum plate. Functioning of the interferometer was therefore verified by visual evaluation of fringe patterns produced by a Na lamp and a laser before and after cycling the mirror transport mechanism at liquid-helium temperatures. Room temperature interferograms were, however, obtained by using a thermocouple and various filters. An unfiltered raw interferogram made over a 2-cm scan is shown in Figure 28-5. The slant of the interferogram indicates a change in modulation efficiency over the scan.

Figure 28-5. Interferogram of a Thermal Source Obtained With the Interferometer and a Room-temperature Detector

28–5 COMMENTS

A liquid-helium-cooled Michelson interferometer has been constructed and tested. Solutions to the three problem areas of beam splitters, mirror drive and lubrication, and interferometer adjustment and alignment, have been achieved. Development of a flight instrument is considered feasible.

Acknowledgments

The authors wish to acknowledge the contributions of Charles Ellis to this program.

References

1. Spitzer, L., Dressler, K., and Upton, W. L. (1964) *Pub. Astron. Soc. Pacific* **76**: 387.

2. Takayanagi, K., and Nishimura, S. (1961) *Rpt. Inst. Space Res., Japan* **15**: 81.

3. Gould, R. J. (1964) *Annales D'Astrophysique* **27**: 815.

4. Field, G. B., Rather, J. D. G., Aanstead, P. A., and Orszag, S. A. (1968) *Astrophys. J.* **151**: 953.

5. Delmer, T. N., Gould, R. J., and Ramsay, W. (1967) *Astrophys. J.* **149**: 495.

6. Gillett, F. C., and Stein, W. A. (1969) *Astrophys. J. (Letters)* **155** (Part 2): L97.

7. Petrosian, V., and Bahcall, J. (1969) *Astrophys. J. (Letters)* **155** (Part 2): L57.

8. Putley, E. H. (1964) *Infrared Phys.* **4**: 1–8.

9. Jacquinot, P., and Girard, A. (1967) *Advanced Optical Techniques*, A. C. S. van Heel, Ed., North-Holland, Amsterdam, pp. 71–123.

10. Mack, J. E., McNutt, D. P., Roesler, F. L., and Chabbal, R. (1963) *Appl. Opt.* **2**: 873.

11. Heinrich, P. L., Bastien, R. C., Dos Santos, A., and Ostrelich, M. (1966) Engrg. Rpt. No. 8421, Perkin Elmer Corp., Norwalk, Conn.

12. Lewis, E. A., and Casey, J. P. (1952) *J. Appl. Phys.* **23**: 605.

13. Casey, J. P., and Lewis, E. A. (1952) *JOSA* **42**: 971.

14. Renk, K. F., and Genzel, F. (1962) *Appl. Opt.* **1**: 643.

15. Vogel, P., and Genzel, L. (1964) *Infrared Phys.* **4**: 257.

16. Richards, P. L. (1964) *JOSA* **54**: 1474.

17. Born, M., and Wolf, E. (1964) *Principles of Optics*, ed. 2 rev., Macmillan, p. 40.

18. Air Force Materials Laboratory (1968) *Cryogenic Materials Data Handbook*, Vol. 2 rev., Air Force Systems Command, Wright-Patterson AFB, Ohio.

19. Vanasse, G. A., and Sakai, H. (1967) in *Progress in Optics*, Vol. 6, E. Wolf, Ed., North-Holland, p. 280.

20. Foskett, C. T., and Weinberg, J. M. (1969) *Appl. Opt.* **8**: 2185.

29. A Half-Wavenumber Michelson Interferometer Operating at Cryogenic Temperatures for the Spectral Region of 5 to 15 μm

James Engel and Geert Wijntjes
Block Engineering, Inc.
Cambridge, Massachusetts

Andrew Potter
Manned Spacecraft Center
NASA
Houston, Texas

Abstract

Mechanical, cryogenic, optical, and electronic design details of a high-resolution Michelson interferometer constructed for operation at cryogenic temperatures are presented. The merits and problems of multiplex spectroscopy under reduced background flux conditions and the resulting enhanced detectivity are reviewed. Anticipated sensitivities to extended sources and astronomical point sources when the interferometer is equipped with a cooled telescope are also discussed.
This work was supported by NASA under Contract No. NAS9-9807.

It is well known that extrinsic photoconductors like Ge(Hg) and Ge(Cu), as well as HgCdTe detectors, when cooled to low enough temperatures become almost wholly photon flux noise-limited (Blip detectors). An analog of the situation that occurs in the visible spectrum (the multiplex advantage of the interferometer spectrometer is in certain instances canceled out by the increased photon flux noise) might therefore be expected.

A closer look at the behavior of these detectors under low photon flux conditions does not bear this out. Contrary to the case of photomultipliers in the visible spectrum, these long-wavelength photodetectors do not possess a reasonably noise-free internal gain mechanism. Under low-background conditions, we are soon confronted with the problem that further reduction of the background photon flux does not correspondingly improve the detectivity of the detector and must deal with other limiting noise sources such as preamplifiers, vibration, microphonics, and boil-off. In such case, of course, the multiplex advantage of $\sqrt{N}$, where N is the number of spectral resolution elements observed, does apply.

But there is also an additional factor: dielectric relaxation. This phenomenon causes the detector time constant to become very long for low-background

flux conditions. This dielectric relaxation time constant is inversely proportional to the photon flux incident on the detector. When the detector is operated at frequencies in excess of its cutoff frequency, the system gain—and hence the signal-to-noise ratio—is therefore directly proportional to the photon flux on the detector. It has been suggested that a fast response-time detector can be obtained by adding a certain photon flux to achieve the required response time. A multiplex spectrometer does this inherently, of course, and therefore we find that if the system is noise-limited the interferometer not only has the usual $\sqrt{N}$ gain in signal-to-noise, but it can have a gain of as much $N\sqrt{N}$. That this situation is real can be seen from Table 29-1.

Table 29-1. Parameters of He:Hg Detector and Preamplifier (77°K Blackbody Background)

Photoconductive gain (at 10-μm optimum bias)	10 mA/W
Resistance	$10^{10}\ \Omega$
Capacitance	2 pF
Time constant	20 msec
Background-noise-limited NEP	2×10^{-14} W/Hz$^{\frac{1}{2}}$
Equivalent amplifier noise current (at 200 Hz)	10^{-16} A/Hz$^{\frac{1}{2}}$
System-noise-limited NEP (at 200 Hz)	10^{-13} W/Hz$^{\frac{1}{2}}$

The cooled interferometer we used is basically the standard Block Engineering Model 296 high-resolution Michelson interferometer modified to provide for remote alignment of the main interferometer and a reference system free from realignment problems. The gas bearing that serves as the support for the moving mirror was not modified because it performs well at both room temperature and 77°K. (See Figure 29-1.)

The double-walled construction of the instrument provides a reservoir for the cryogen. The gas produced by the boiloff is collected near the top of the instrument and fed into the gas bearing by means of a flexible bellows arrangement. The inlet pressure on the gas bearing is controlled by means of a pressure-regulated heater located in the cryogen to increase or decrease the boiloff rate. Spent gas is released from the instrument to ambient by means of a one-way pressure relief valve or alternatively fed into a series of valves and ports designed to prevent frosting of the

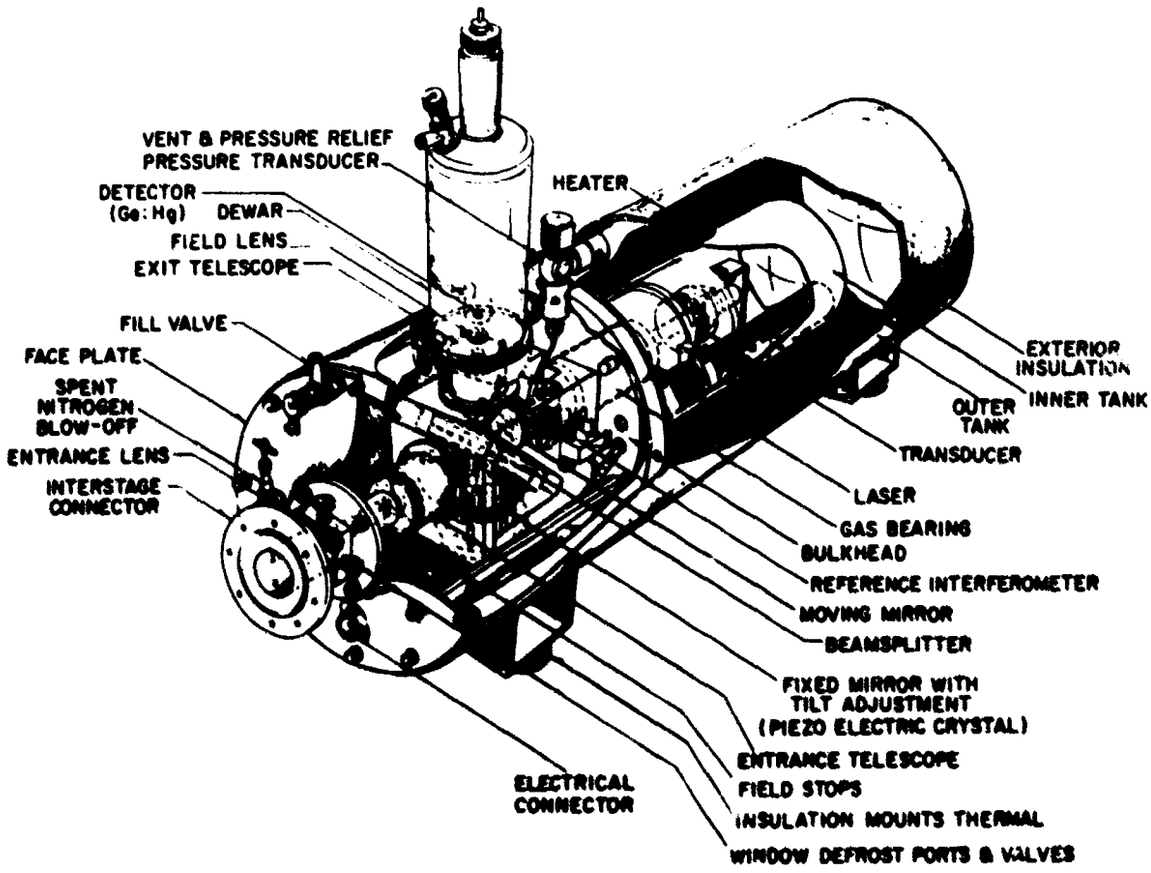


Figure 29-1. Half-Wavenumber Michelson Interferometer Modified for Operation at Cryogenic Temperatures

entrance window. Since the gas consumption of the bearing is relatively high, the insulation requirements for the dewar are minimal and the insulation provided serves mainly to prevent frost buildup and protect personnel.

An independent reference interferometer built into the instrument behind the moving mirror provides both white-light and laser reference signals. The use of retroreflectors in this system eliminates the need for realignment after initial adjustment at room temperature since the entire unit, including white light and laser, will operate at the cryogen temperature.

Adjustment of the fixed mirror in the main interferometer is provided for by a combination of mechanical and piezoelectric elements. Initial alignment at room temperature is accomplished by using the Model 296 wedge gear arrangement. Alignment is maintained through the cooling cycle by varying the voltage applied to piezoelectric elements located under the fixed mirror. These elements provide for tilt adjustment within 1 min of arc.

The optical system is throughput-matched to 1 cm^{-1} at 5 μm. It consists of a simple objective entrance lens and field stop, two telescopes of the Cassegrain type as input and output devices for the interferometer cube, an antireflection-coated germanium beam splitter, a field lens, filter holder, and detector. The optical elements are all at ~77°K except for the filter and detector, which are at ~4°K. Baffling will be used to provide the system with good off-axis rejection capabilities. Initial testing is being done with liquid nitrogen as the main cryogen, but it should be possible (although perhaps expensive) to operate with liquid neon in the entire system.

The electronics are straightforward. The preamplifier is used in the operational amplifier mode, cooled to liquid nitrogen temperature, with the amplifier feedback resistor cooled to liquid helium temperature. Provision is made to monitor the resistance of the detector. An InAs photon source located in the detector dewar controls detector gain and frequency response.

This instrument will initially be used with a cold-gas reaction cell for measuring simulated airglow phenomena. Table 29-2 lists collected specifications for the cooled interferometer and Table 29-3 lists the estimated sensitivity of the instrument to extended sources and stellar sources.

Table 29-2. Collected Specifications for Cooled Interferometer

Spectral range	5 to 15 μm
Wavelength precision	0.1 cm^{-1}
Resolution	0.5 cm^{-1} maximum
Optical retardation	0.25, 0.5, 1.0, 2.0 cm
Scan velocity	0.63 cm/sec
Velocity error	<0.5 percent
Sampling interval	0.6328 μm $\times \frac{1}{2}$, 1, 2, 4, 8, 16
Sampling precision	<0.01 μm
Detector	Ge:Hg (T~4°K)
Field of view	7° full angle (1.2×10^{-2} ster)
Entrance aperture	2.6 (1.8-cm diam.)
Beam splitter	antireflection-coated germanium
Instrument operating temperature	~78°K (LN_2)
Hold time	6 hr
Operating altitude	zero to 140,000 ft
Instrument size	36×14×28 in.
Instrument weight	optical head, 125 lb; control unit, 50 lb

Table 29-3. Estimated Sensitivy of Cooled Interferometer

Sensitivity to Extended Sources	
Observation time	1 min
Resolution	1 cm^{-1}
Background	77°K blackbody
NESR, at 10 μm	10^{-12} W/cm^2-sr-cm^{-1}
Sensitivity to Stellar Sources	
Observation time	10 min
Resolution	1 cm^{-1}
Field of view	2 mrad
Altitude	100,000 ft
Telescope aperture	10-in. diam.
Telescope temperature	200°K
Emissivity of optics	<1 percent
NESI, at 10 μm	2×10^{-17} W/cm^2-cm^{-1}

Discussion

Q. (P. L. Richards, University of California): Could you repeat your statement about the multiplex advantage?

A. (G. Wijntjes): By changing the wording from multiplex advantage to observation time advantage it is clear that if the spectrometer covers m spectral elements but there are only n elements actually present, the signal-to-noise advantage will vary as $\sqrt{m/n}$.

Contents

30. A Large-Aperture Field-Widened Interferometer-Spectrometer for Airglow Studies

Alvin M. Despain, Fon R. Brown, Jr., Allan J. Steed, and Doran J. Baker
Electro-Dynamics Laboratories
Utah State University

Abstract

A field-widened interferometer-spectrometer with a 5-cm aperture has been created and applied to airglow studies. The instrument is capable of a spectral resolution of 1 cm^{-1} and a field of view to over 5°, resulting in a field-widened throughput gain of 19 compared with a standard Michelson interferometer of the same aperture. The spectral response is 0.4 to 3 μm. The digital recording technique that is used compensates for source variations.

30–1 INTRODUCTION

A large-aperture field-widened interferometer-spectrometer system has been designed and developed by the Electro-Dynamics Laboratories of Utah State University.* The instrument and its data-handling system are specifically intended for the study of aeronomy, where the primary emissions are extended sources.

The light-gathering capability of a Michelson interferometer is proportional to the square of its acceptance angle. Accordingly, several approaches to increasing the field of view of an interferometer have been proposed. The synchronous wedge and drive mirror technique suggested by Boucharcine and Connes[1] was used to provide field compensation in the interferometer reported herein. The salient features of the design are: (1) field of view widened to 5° full angle, (2) a 5-cm clear aperture, (3) a V-groove in granite translational drive with 5-cm drive capability, (4) intensity compensation for source variations, and (5) direct digital recording.

* This work was supported by a grant from the National Science Foundation to Utah State University.

30–2 AIRGLOW SURVEYS IN THE NEAR INFRARED REGION

Chamberlain[2] and Broadfoot and Kendall[3] have published very useful spectra of the night airglow, valuable for identifying the emitting constituents of the upper atmosphere and evaluating the relative importance of the various processes that lead to such emissions. Their work, however, is presented in detail out to only 1 μm. The USU field-widened interferometer-spectrometer is being used to extend the survey to longer wavelengths, greater wavelength resolution, and greater time resolution, and to overlap into a portion of the visible spectrum in order to provide a comparison and detect features that may have been missed.

Although the interferometer spectrometer can be used at wavelengths as short as ultraviolet, its primary utility is in the near infrared, where the Fellgett

multiplex advantage[4] is fully realized and where no high-resolution and only a little moderate-resolution survey work has been done. At wavelengths beyond about 3 μm, thermal emissions from the lower atmosphere mask the airglow from ground observation. Major emphasis will therefore be on airglow emissions in the spectral range from about 0.6 through 3 μm, with maximum resolution of about 1 cm^{-1}. These measurements will directly help to answer many of the leading questions about upper atmospheric photochemical reaction mechanisms.

The hydroxyl gas emissions (Meinel[5]), as a group the brightest of all night airglow emissions, comprise a complex spectrum. High-resolution measurements of the OH spectrum are needed, as a function of both altitude and solar zenith angle, in order to understand upper atmospheric chemical kinetics. Vibrationally-excited $OH^{\ddagger}$ can possibly be created in at least nine different ways, but the primary reaction is probably the one proposed by Bates and Nicolet:[6]

$$O_3+H\rightarrow OH^{\ddagger}+O_2.$$

The $OH^{\ddagger}$ generated by each reaction is, in general, distinct since each reaction is uniquely exothermic or endothermic. Further, several of the reactions are very temperature-dependent. The more detailed spectroscopic information now being obtained with the interferometer-spectrometer can be used to correlate with the recently developed theoretical models (Hampson[7]) for the chemical kinetics of the upper atmosphere.

The radiations from the O_2 molecular electronic states $O_2(^1\Delta_g)$ and $O_2(^1\Sigma_g)$ will also be monitored throughout the night and as far as possible into the twilight on either side of night (Megill et al. [8]). The realization is growing that the $O_2(^1\Delta_g)$ species plays a significant role in the photochemistry of the upper atmosphere (Gattinger and Vallance-Jones[9]). It can be formed from the photodissociation of ozone, as

$$O_3+h\nu_{uv}\rightarrow O_2(^1\Delta_g)+O(^1D).$$

Then, since $O_2(^1\Delta_g)$ has 0.98 eV of energy, the important detachment reaction

$$O_2^-+O_2(^1\Delta_g)\rightarrow 2O_2+e^-$$

can occur. Nitric oxide, with its low ionization potential, may also be generated by

$$O_2(^1\Delta_g)+N\rightarrow NO+O,$$

which is important to D-region ionization. An alternative source, of perhaps even greater importance, is $O_2(^1\Delta_g)$ itself, which will ionize at significantly longer wavelengths than the ground state O_2. Absolute photon emission rates for the various bands of molecular oxygen need to be established.

Also being initiated with the interferometer-spectrometer is an experimental study of the Meinel bands of N_2^+ in the near infrared region, made feasible by the high sensitivity-resolution trade-off capability of the instrument. It should be possible to obtain measurements into the twilight. The state of interest is the ($^2\Pi$, $v=1$). It is probably produced by the charge transfer reaction (Omholt,[10] Hunten,[11]):

$$O^+(^2D)+N_2\rightarrow N_2^+(^2\Pi)+O.$$

This reaction is thought to be the significant one because perhaps 30 percent of the O^+ is excited by solar radiation to the long-lived ^2D state. What we expect to see is the anomalous distribution of the vibrational bands, which has previously led to the deduction of incorrect vibrational temperatures. The expected measurements will give us a partial look at the $O^+(^2D)$ production and at the above charge transfer reaction, which is an order of magnitude faster than the familiar

$$O^++N_2\rightarrow NO^++N.$$

Lastly, O^{++} has a resonance line at 5007 Å, and it should be possible to learn something of the amount of O^{++} in the atmosphere by measuring the intensity of this line during twilight. Nakada and Singer[12] predict that large amounts of doubly ionized oxygen should be present at heights on the order of one earth radius, especially over the equator. They predict an intensity of about 50 rayleighs, but it must certainly be weaker than this or it would have been discovered by now. To be able to detect the 5007 Å line in the presence of OH bands in this region of the spectrum, it is necessary to have a spectrum-scanning device such as the field-widened interferometer-spectrometer.

30-3 THEORY AND DESIGN CONSIDERATIONS

30-3.1 Night Airglow Radiance Levels

The measurement of twilight and night airglow requires a system limited only by the sky background radiation. For a spectral measurement of the night airglow, this background is generally assumed to be the airglow continuum, which has a magnitude of approximately 1 rayleigh per angstrom in the visible portion of the spectrum (Chamberlain[2]).* To effectively measure spectra at this radiance level requires a large throughput optical system such as a field-widened interferometer. A sketch of a wedge-end-mirror field-widened interferometer is shown in Figure 30-1. Other methods of field widening are known, but this is the only configuration that will be discussed in this paper.

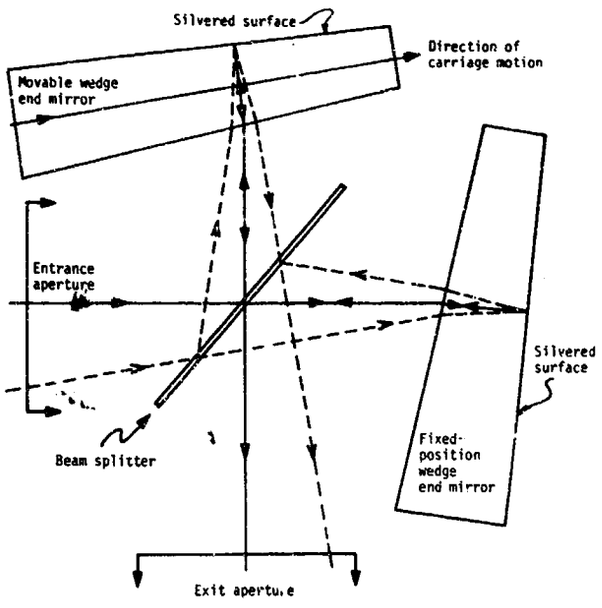


Figure 30-1. Sketch of Wedge End-Mirror Field-Widened Interferometer Optics

30-3.2 Field-Widening

An optical throughput of about 1 cm^2 sr is needed to match contemporary detectors. This requirement, along with the desire for good resolution (1 cm^{-1}), led to the design of a wide-aperture field-widened interferometer configured after a Michelson but modified to increase its optical input acceptance angle and in turn its sensitivity to extended sources. In the usual Michelson interferometer the path-length difference depends on the angle at which radiation enters the instrument. By field-widening we attempt to make the path difference invariant with the entrance angle.

* The conversion between photon emission rate R in rayleighs and radiance N in W cm^{-2} $ster^{-1}$ is: $R = 2\pi\lambda N \times 10^{13}$, where λ is the wavelength in micrometers.

A technique suggested by Bouchareine and Connes[1] accomplishes such field compensation to the first order by means of prisms in place of the usual mirrors in the Michelson interferometer (Figure 30-1). The back sides of the prisms are silvered, and the optical system is aligned as in a typical interferometer. Instead of obtaining differential path-length changes by moving one mirror parallel with the rays of light that impinge on it, in this field-widening technique a prism is moved parallel with its 'apparent mirror position,' inserting more optical material in one beam in order to achieve an effective differential path-length increase. 'Apparent mirror position' refers to the apparent position of the coated back surface of the prism as it appears to an observer in front of the prism, looking through the optical material. The invariance with incident angle of the differential path length can be shown to follow from Fermat's principle.

To achieve the field-widening increase in optical throughput, large prisms must be constructed, and displaced during operation, so that wavefront distortions are less than $\lambda/10$ across 10 cm. This requires extreme optical homogeneity and mechanical positioning control. The range of wavelengths that can be examined is limited by variations in the transparency and refractive index that occur as a function of the incident radiation wavelength.

30-3.3 Interferometer Resolution

The resolution capability of a Michelson interferometer is generally taken to be

$$\Delta\bar{\nu} = \frac{1}{\delta}, \tag{30-1}$$

where

$$\Delta\bar{\nu} = \bar{\nu}_2 - \bar{\nu}_1 \tag{30-2}$$

is the difference in wavenumber between two resolvable adjacent monochromatic lines and δ is the optical path difference. This criterion for resolution will be used throughout this chapter since it applies equally well to either a standard Michelson or a field-widened Michelson interferometer.

The retardation of the field-widened interferometer, however, is calculated differently from that of the standard Michelson interferometer. In the Michelson there is a simple relationship between the displacement of the mirror and the optical path difference, namely,

$$\delta = 2d, \tag{30-3}$$

where d is the interferometer end-mirror displacement. The relationship between the optical path difference and distance of motion of the wedge prism along its apparent mirror image plane in the field-widened system (see optical block diagram in Figure 30-1), although not so simple, can nevertheless be determined. In Figure 30-2 the field-widening prism is

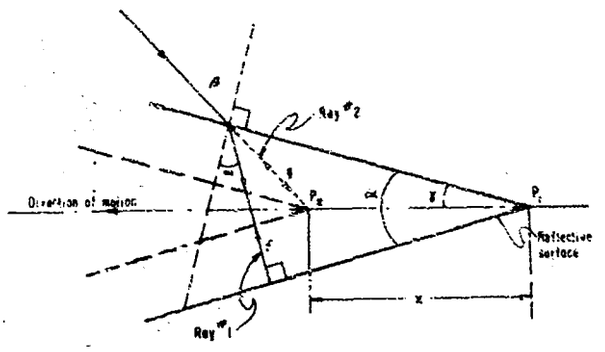


Figure 30-2. Field-Widened Interferometer Wedge End-Mirror Geometry

pictured near the apex, where the wedge is moved a distance x from P_1 to P_2. The optical path difference can be expressed by

$$\delta = 2(nf - g), \tag{30-4}$$

where n is the index of refraction of the prism. From the geometry of Figure 30-2 the optical path difference can be expressed as

$$\delta = 2x \left\{ n \sin \alpha + \tan \gamma \, \frac{n^2 \sin^2 \alpha - 1}{\cos \beta} \right\} \cos \gamma, \tag{30-5}$$

where n is the index of refraction, α is the prism angle, $\beta = \sin^{-1}(n \sin \alpha)$, and $\gamma = \tan^{-1}(\tan^2 \alpha / \tan \beta)$. For the values $\alpha = 8°$ and $n = 1.5$,

$$\delta = 0.21x. \tag{30-6}$$

Thus, to obtain a resolution of 1 cm^{-1}, a drive length of at least 4.7 cm is required.

30–3.4 The Field of View

The field-of-view limitation for the Michelson interferometer with field compensation has been derived by Bouchareine and Connes[4] for the usual three types of aberrations: spherical, chromatic, and astigmatic. The spherical aberration limit (full angle) is

$$\theta_s^2 = 8n\sqrt{\lambda/d}, \tag{30-7}$$

(compared with $\theta_s^2 = 4\lambda/d$ for the standard Michelson), where d is the effective retardation of the mirror and λ is the wavelength in micrometers. The chromatic aberration limit is

$$\theta_c^2 = \frac{4n\lambda(n^2-1)}{\Delta n d}, \tag{30-8}$$

where n is the arithmetic mean between n_1 and n_2 and $\Delta n = n_1 - n_2$. The astigmatic aberration limit exhibited by the field-widened interferometer is

$$\theta_a^2 = \frac{8\lambda}{d \tan^2 \alpha}, \tag{30-9}$$

where the computation has been made for the greatest astigmatic effect, namely, in a plane perpendicular to the surface and extending from the narrow to the thick end of the prism and, again, α is the prism angle.

Equations (30-7), (30-8), and (30-9) yield the theoretical field of view (FOV) of the field-widened interferometer. The aberrations will impose a limit on the field of view, depending of course on the wavelength range, resolution, prism material, and prism angle. For the airglow measurements, the required system parameters resulted in a system whose limit is chromatic aberration. This is the case even though fused silica of very low dispersion was chosen as the optical material for the beam splitter and prisms. The computed theoretical field-of-view limit is 9.7° full angle. It has been verified experimentally that excellent results are obtained at 5°. With further effort we expect to attain field compensation to the theoretical limit.

30–3.5 Drive and Carrier Configuration

The design of a drive mechanism and mirror-carrier configuration for a field-widened interferometer

presents a difficult problem. The end mirror must be held in alignment to an extremely close tolerance. If the contrast function of the instrument is to be maintained, then the wavefront distortion of the two legs of the interferometer must be less than $\lambda/4$ when they recombine at the beam splitter. This condition must hold for any position of the driven end mirror. The mirror must then move in translation without rotating more than

$$\phi = \lambda/4D, \tag{30-10}$$

where D is the diameter of the end mirror. For a 5-cm system at 4000 Å, ϕ is only 2×10^{-7} rad.

The mirror carrier system used by Otteson[13] for his bent-crystal gamma-ray spectrometer appeared to be the most acceptable design approach. His system was checked with a Twyman-Green interferometer and was found to have repeatable rotational position stability, when driven in one direction, to less than 1×10^{-7} rad over 15 cm of the drive. Several improvements suggested by Otteson[13] (private communication) were later incorporated into the interferometer drive and carrier design. The resulting ways and carrier configuration also exhibited rotational stability to less than 1×10^{-7} rad over 10 cm of drive. The exact stability was beyond the resolution capability of the measuring instrument.

30-3.6 Measurement Time

The first measurements were undertaken in the visible and near infrared with a multiplier phototube used as the detector. It can be shown that the total time T required to make one interferometer scan to achieve a desired signal-to-noise ratio (s/n) in the resulting optical spectrum is

$$T = \frac{2\pi eG[\sigma/(\sigma-1)](I_{dk}+I_s)}{(NA\Omega ER)^2}\,(s/n)^2, \tag{30-11}$$

where $e=1.6\times10^{-19}$ C, G is the gain of the photomultiplier, σ is the dynode gain, I_{dk} and I_s are respectively the dark and signal currents (amperes) at the anode, N is the radiance (W cm^{-2} sr^{-1}) of a spatially extended monochromatic source, $A\Omega$ is the optical throughput (cm^2 sr) and E is the optical efficiency. The anode radiant sensitivity or responsivity (A/W) is

$$R = \frac{eG}{h\nu}Q_e, \tag{30-12}$$

where h_ν is the photon energy (W-sec) and Q_e is the quantum efficiency (photoelectrons/photon).

When the signal current is much greater than the dark current $I_s \gg I_{dk}$,

$$I_s = N_{bw}A\Omega ER, \tag{30-13}$$

and Eq. (30-11) reduces to

$$T = \frac{2\pi N_{bw}eG(\sigma/\sigma-1)}{N^2A\Omega ER}\,(s/n)^2, \tag{30-14}$$

where N is the radiance of the monochromatic signal source and N_{bw} is the total radiance in the instrument bandpass.

The interferometer scan time required to obtain a spectrum, say with an (s/n) of unity at 0.5 μm and a resolution of 1 cm^{-1}, of an extended source whose intensity is 1 rayleigh/Å (with $N_\lambda = 3\times10^{-14}$ W cm^{-2} sr^{-1} μm^{-1}) between 0.4 and 0.7 μm is computed as follows. From Eq. (30-6) the drive length has already been computed to be 4.7 cm. On the basis of the parameters for an EMI S-20 photomultiplier and an interferometer throughput $A\Omega$ of 0.1 cm^2 ster, Eq. (30-14) gives the calculated time required to take the interferogram as 11 min.

The ratio of the optical throughput of the field-widened interferometer to that of a standard Michelson, where both have collectors of the same diameter, is given by:

$$\frac{A\Omega_{FW}}{A\Omega_M} = \frac{\theta_{FW}^2}{\theta_M^2} = \frac{\bar{\nu}}{8\Delta\bar{\nu}}\,\theta_{FW}^2, \tag{30-15}$$

where θ_{FW} and θ_M are the respective full-angle fields of view. Consequently, the relative throughput of the interferometer without field-of-view compensation at the same wavelength and resolution is only about 5 percent, namely, 5×10^{-3} cm^2 sr. The time required to make a scan and achieve the same (s/n) in the resulting spectrum is about 210 min, whereas the field-widened interferometer requires only 11 min.

When a photoconductive detector is used, the formula for computing the time required to make one

interferometer scan and achieve a specified (s/n) in the resulting spectrum is

$$T=\frac{\pi A_d}{(NA\Omega ED^*)^2}(s/n)^2, \qquad (30\text{-}16)$$

where A_d is the area of the detector.

Consider the time required to make an interferogram of an extended source whose spectral radiance is 1×10^{-8} W cm^{-2} sr^{-1} μm^{-1} at 2.5 μm. At this wavelength the instrument throughput is about 0.4 cm^2 ster. A 1-cm-diam. cooled (195° K) lead sulfide detector is used. The scan time to obtain an optical spectrum with an (s/n) of unity at a resolution of 1 cm^{-1} (6 Å) is about 30 sec. The time required by a standard Michelson interferometer with a collector of the same diameter would be about 2 hr.

30-4 INSTRUMENTATION SYSTEM

30-4.1 Configuration

The complete interferometer-spectrometer system is shown in Figure 30-3. The design goal was to achieve a clear aperture of 5 cm and maximize the field of view while maintaining a 1-cm^{-1} resolution. A 7.6-cm beam splitter and wedge mirrors, having a back surface of 10-cm diam. and wedge angles of 8°, all configured to maintain a wavefront distortion within λ 10, were specially ground. The beam splitter consists of a thin layer of silver on the interface between two fused silica prisms that together form a cube. The low dispersion of quartz helps to minimize the chromatic aberration, which is the limiting parameter on the widened field. The optical components of the interferometer are pictured in Figure 30-4.

Figure 30-3. Complete Interferometer Installation

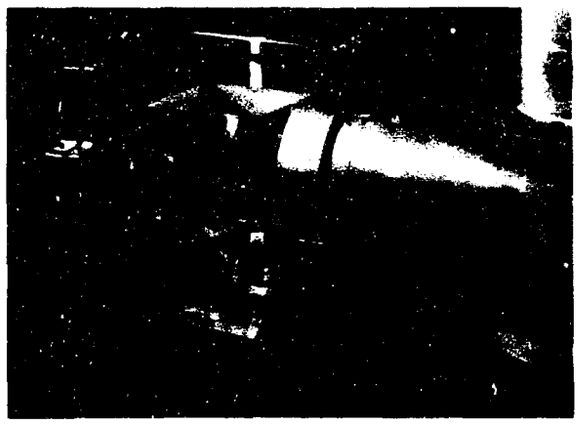


Figure 30-4. Field-widened Beam Splitter and Wedge End Mirrors

The base of the interferometer was constructed of black granite with a V groove, following the technique used for the bent-crystal gamma-ray spectrometer. The surface flatness on the top and in the groove, where the end mirror carrier was indexed, was maintained at 25×10^{-6} in. (Figure 30-5).

The optical carrier was constructed of invar stainless steel, whose very low expansion coefficient is near that of granite and quartz. To minimize the torque

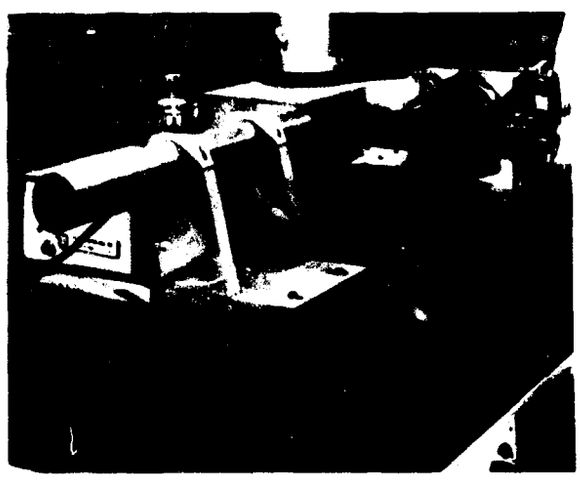


Figure 30-5. Granite Block and End-Mirror Carrier

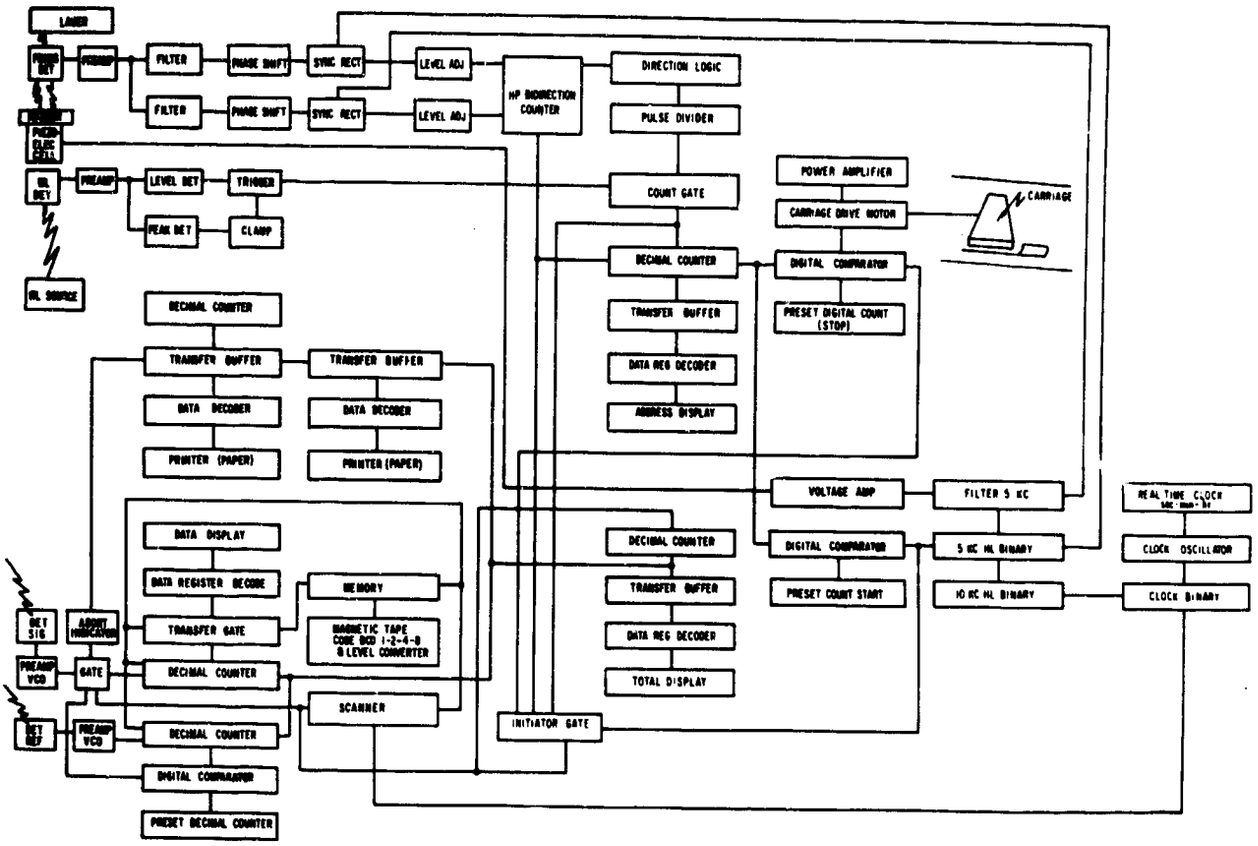


Figure 30-6. Block Diagram of Field-widened Interferometer Electronics

moments the drive was situated to bear colinearly with the center of mass of the carrier. The reference and indexing buttons were extended sufficiently to maintain the rotation of the driven mirror to less than 10^{-7} rad. The block and carrier are also shown in Figure 30-5.

The signal control and monitoring electronic systems are shown in block diagram form in Figure 30-6. The outputs from the signal and background detectors are amplified and then fed to voltage-controlled oscillators (VCO). The output of each VCO is fed into a digital counter; the reference counter gates the signal counter so that the count in the latter is normalised to the background and the background fluctuations are removed from the data. The count in the signal counter is then transferred to magnetic tape on command from the zero crossing of the laser reference.

Figure 30-7 is a photograph of the interferometer electronics console.

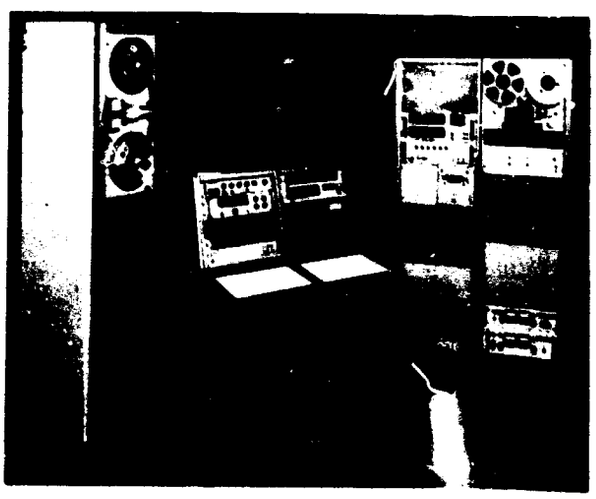


Figure 30-7. Interferometer Electronics Console

30-4.2 Performance

A raw interferogram of the night sky on 23 February 1970, taken in 1.5 min, is shown in Figure 30-8. The computer-performed Fourier transform of this interferogram is shown in Figure 30-9. These early data were processed to a resolution of only $10\ \mathrm{cm}^{-1}$. In our present night-sky measurements program, we are attempting to obtain spectra fully resolved to $1\ \mathrm{cm}^{-1}$.

Unlike those of the sustained airglow survey scheduled to be conducted, these early measurements were made from Logan City; street light scattered from the atmosphere above the observation dome is therefore very much in evidence. The 5460-, 5770-, and 5790-Å mercury lines from these street lights, along with the 5577-Å atmospheric oxygen green line, make up the four dominant lines seen in Figure 30-9. Many other atmospheric emission lines such as the 6300-Å oxygen red line are also prevalent.

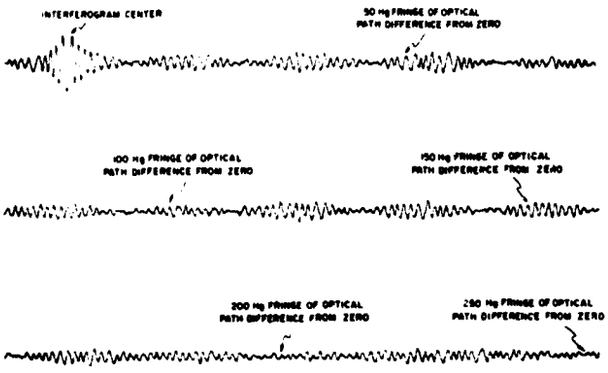


Figure 30-8. Night Airglow Interferogram

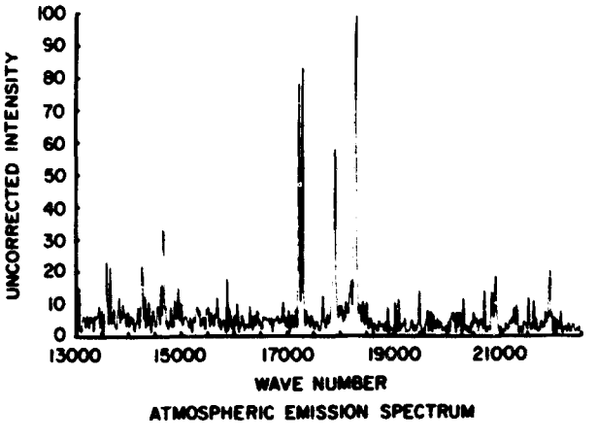


Figure 30-9. Night Airglow Spectrum

References

1. Bouchareine, P., and Connes, P. (1963) *J. Phys. Radium* **24**: 134.

2. Chamberlain, J. W. (1961) *Physics of the Aurora and the Airglow*, Academic Press.

3. Broadfoot, A. L., and Kendall, K. R. (1968) *J. Geophys. Res.* **73**: 426–429.

4. Stewart, S. P. (1966) Ph. D. Dissertation, Utah State University.

5. Meinel, A. G. (1950) *Astrophys. J.* 111: 555.

6. Bates, D. R., and Nicolet, M. (1950) *J. Geophys. Res.* **55**: 301–327.

7. Hampson, J. (1965) CARDE Rpt, CARDE TM 1690165.

8. Megill, L. R., Baker, K. D., Despain, A. M., and Baker, D. J. (1969) Am. Geophys. Union, National Fall Meeting, San Francisco, 18 December 1969.

9. Gattinger, R. L., and Vallance-Jones, A. (1966) *Planetary Space Sci.* **14**: 1.

10. Omholt, A. (1957) *J. Atmospheric Terrest. Phys.* **10**: 320.

11. Hunten, D. M. (1958) *Ann. Geophys.* **14**: 167.

12. Nakada, M. P., and Singer, S. F. (1968) *Geophys. J.* **15**: 163–181.

13. Otteson, H. (1967) Ph. D. Dissertation, Utah State University.

Discussion

Q. (P. Feldman, Johns Hopkins University): What is the sensitivity of your instrument in rayleighs?

A. (F. Brown, Jr.): To see a 1-rayleigh 1-Å-wide line source at 5000 Å with a signal-to-noise ratio of unity in the transformed interferogram takes a scan of less than 1 min.

Contents

31. Application of Fourier Spectroscopy Technique to the Study of Relaxation Phenomena

Randall E. Murphy and Hajime Sakai
Air Force Cambridge Research Laboratories
Bedford, Massachusetts

Abstract

A technique that exploits the multiplex advantage of Fourier spectroscopy can be used for studying time-varying sources. The time behavior of each spectral element is observed simultaneously with that of all the others. The technique is demonstrated by an experimental measurement of the relaxation of the hot bands of CO_2 and N_2O.

31–1 INTRODUCTION

The advantages that can be realized by using Fourier spectroscopy techniques have been demonstrated in several types of spectroscopic measurements.[1–4] Such spectroscopic measurements have thus far been limited to the observation of sources that are stationary in time—or at least stationary by comparison with the total observation time. In many instances the radiation history of time-varying sources is of primary interest. For example, time-resolved spectroscopy is concerned with the measurement of radiative lifetimes, energy exchange, quenching, and relaxation processes in general.

The conventional approach of a time-resolved spectroscopic study requires the use of classical grating spectrometers or spectral bandpass filters, in which case a set of spectral lines must be observed sequentially in order to determine the time behavior of the entire spectrum. By contrast, the multiplex technique of Fourier spectroscopy provides for simultaneous observation of all spectral elements, and thereby yields a significant gain in signal-to-noise ratio for the same measurement time.

A technique that exploits the multiplex advantage of Fourier spectroscopy in the study of time-dependent phenomena is presented in this paper. Experimental demonstrations of its application to the relaxation of vibrationally excited CO_2 are described.

31–2 THEORY

This discussion applies to discrete positioning of the movable mirror of a Michelson interferometer, that is, to the step-and-stop mode of operation. The technique can easily be extended to apply to a continuously driven movable mirror by making some rather simple modifications.

Let $B(\sigma,t)$ at time t and wavenumber σ be the spectral density incident on the entrance aperture of the interferometer. The output at the detector is given by

$$F(x,t) = \int B(\sigma,t) \cos 2\pi\sigma x \, d\sigma, \qquad (31\text{-}1)$$

where x is the optical path difference. If the source is stationary, the interferogram signal $F(x,t)$ can be integrated for some preset time to obtain $F(x)$. For time-varying sources its time-dependence can be determined by monitoring the instantaneous value of the interferogram signal.

For a fixed optical path difference the interferogram signal is recorded from initial excitation of the source until some later time that is long by comparison with the overall relaxation time. The signal-to-noise ratio in the interferogram measurement can be improved as desired by inducing intermittent excitation and relaxation of the source. A time series of interferograms is obtained for N observations of $F(x,t)$ by recording the interferogram signal for the same relaxation phenomena at successive optical path differences x_n, where $n=1,\ldots,N$. By applying the Fourier transformation to the interferogram $F(x,t)$, the spectrum $B(\sigma,t)$ at any time t during the relaxation process can be recovered as:

$$B(\sigma,t)=\int F(x,t)\cos 2\pi\sigma x\,dx. \qquad (31\text{-}2)$$

An observation of the relaxation for a spectral component σ can be extracted from $B(\sigma,t)$ from initial excitation until relaxation is complete. Since all spectral elements in the band of interest are observed simultaneously, the multiplex advantage is realized. The extent of the time resolution is governed by the time response of the detector and associated circuitry.

31-3 APPLICATION

Figure 31-1 is a schematic drawing showing the setup of the infrared source chosen to demonstrate the technique. Nitrogen is known to be excited to high vibrational levels by a microwave discharge. The near resonance between the energy levels of N_2 at 2331 cm^{-1} and of CO_2 at 2340 cm^{-1} produces the very efficient transfer of vibrational energy from the N_2 to the CO_2 molecule. The vibrational excitation of CO_2 produced by mixing active N_2 and CO_2 provides an excellent source for observing the decay of vibrationally hot CO_2. The typical vibrational temperatures obtained are in the neighborhood of several thousand degrees Kelvin while the rotational temperature remains relatively cold,[5] and the vibrational hot bands are easily observed.

In our experimental tests, active nitrogen was excited by flowing nitrogen through an Evanson microwave cavity activated by a 120-W 2450-MHz microwave generator. The pressure of N_2 at 1.0 Torr, and of CO_2 at 0.2 Torr, was monitored with an MKS Baratron pressure gauge. The cylindrical reaction cell, 7.5 cm in diameter and 15 cm in length, was evacuated with a 15-cfm fore-pump so that the cell residence time of the gases was approximately 10^{-1} sec. Under these conditions CO_2 should be excited by collisions with N_2, and decay by radiation or deactivation at the walls or through quenching by another molecule. The interferometer that was used has 2.5-cm-diam. optics and a liquid N_2-cooled 4 mm by 4 mm InSb detector.

The detector time response was faster than 4×10^{-5} sec. A block diagram of the recording system is shown in Figure 31-2. The output of the detector

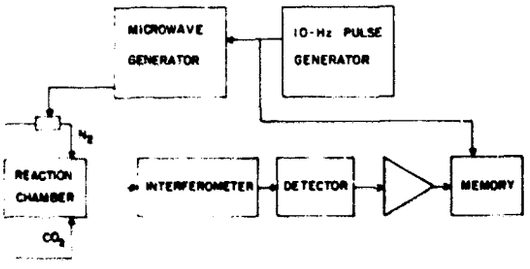


Figure 31-2. Block Diagram of Data Acquisition and Recording

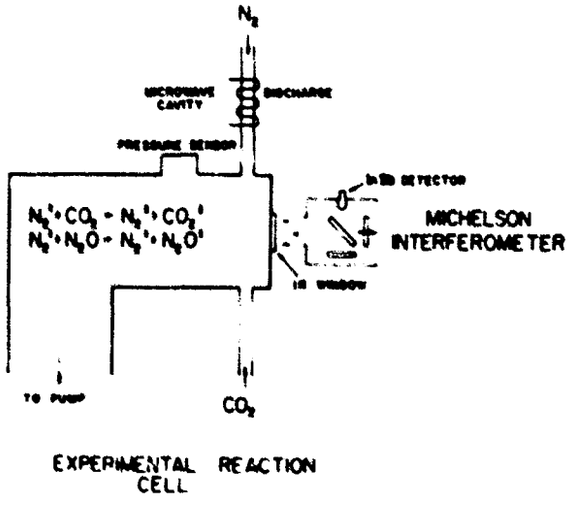


Figure 31-1. Schematic of the Vibraluminescent Experimental Setup (The discharge is created by a 120-W 2450-MHz generator.)

was dc-amplified and registered via a 100-channel Waveform Eductor (built by PAR). An oscillator operating at 10 Hz was used to trigger the microwave generator and provide a temporal reference for the Eductor. Each channel of the Eductor was gated at 10^{-3} sec. A 1-sec integration time in the Eductor provided good signal-to-noise ratios in the interferograms. Figure 31-3 shows the interferograms obtained from four channels of the Eductor. The interferogram $F(x,t)$ was sampled at three times the wavelength of the He-Ne laser reference signal. The four interferograms correspond to the four temporal points. The intensity of the discharge at point 1 is close to the maximum; points 3 and 4 are anomalous features of the N_2 discharge.

Phase-corrected Fourier transforms of the interferograms produced the spectra shown in Figure 31-4.

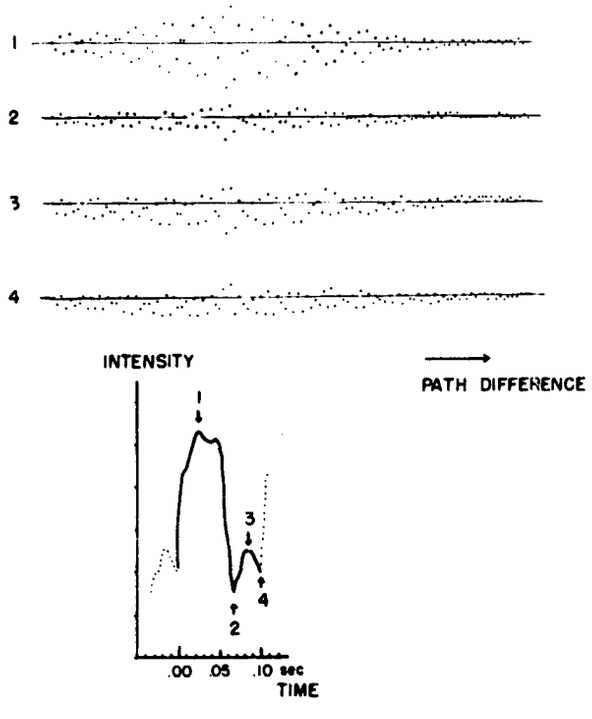


Figure 31-3. Interferograms of the Infrared Source at the Four Times of Interest Indicated on the Waveform (The waveform, which was taken at a particular optical path difference, changes in shape as the optical path is varied.)

The resolution is about 40 cm^{-1} which is insufficient to resolve the vibrational-rotational structure of the bands. Each spectrum shown in Figure 31-5 has been normalized to show the relative shapes of the spectra. When the radiation is at maximum intensity, namely, at point 1, the CO_2 band at 2340 cm^{-1} is much broader than any of the other three observations taken at later times. A spreading of the band to lower wavenumbers is indicated in these observations. This is due to the radiation from the higher vibrational levels of the CO_2 bands, and corresponds to a vibrational temperature of 1500°K. The time difference between positions 3 and 4 is about 0.02 sec; hence, the overall decay of the CO_2 band may be estimated to be approximately 5×10^{-2} sec. This value agrees with other, conventional, measurements we have made with the same gas system.

In addition to the CO_2 bands, the radiation from N_2O at 2223 cm^{-1} was observable at a much lower intensity. Since integrated band intensities of CO_2 and N_2O are known to be comparable, a complexity of the gas kinetics within the reaction vessel is indicated.

It is expected that more refined measurements of time and spectral resolution will lead to more comprehensive analyses, and consequently that the new technique will prove as successful in the study of time-varying phenomena as Fourier spectroscopy has been in the study of stationary sources.

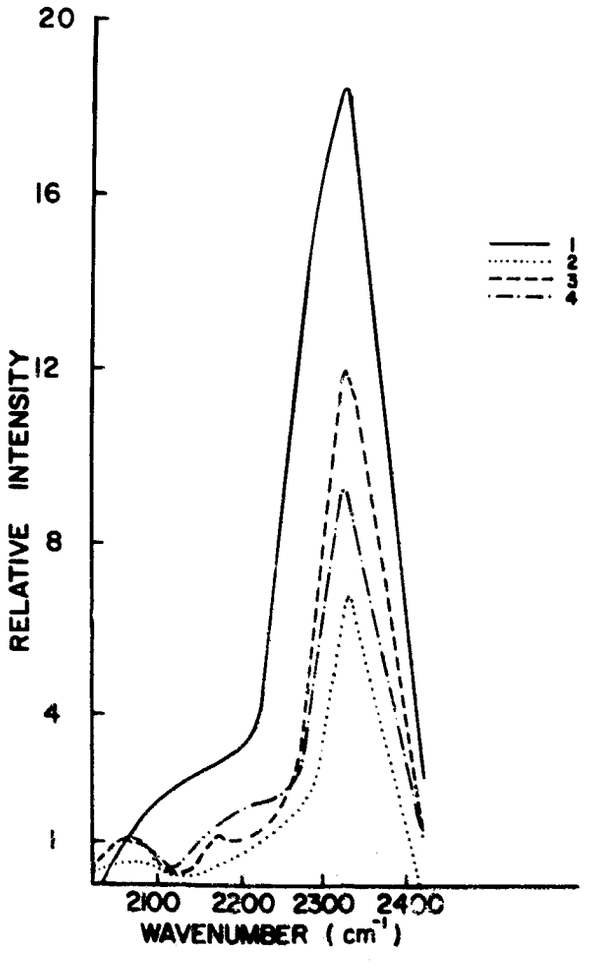


Figure 31-4. Spectral Evolution of the Infrared Source at the Four Selected Times

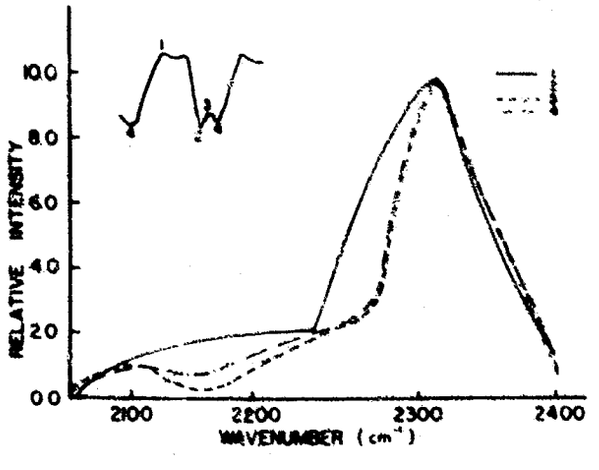


Figure 31-5. Normalized Spectra of the Infrared Source at the Four Times of Interest (The hot bands of CO_2, in evidence at time position 1, are relaxed at the other three time positions. The radiation of N_2O at 2223 cm^{-1} is shown.)

References

1. Connes, J. (1961) *Rev. Opt.* **40**: 45, 116, 171, 231.
2. Richards, P. (1964) *JOSA* **54**: 1474.
3. Connes, J. and Connes, P. (1966) *JOSA* **56**: 896.
4. Hanel, R., et al. (1969) *Appl. Opt.* **8**: 2059.
5. Legay, F. (1964) *J. Phys. Radium* **25**: 999.

Discussion

Q. (B. F. Hochheimer): Did you try relaxation measurements on CO_2 with helium in cell?

A. (R. Murphy): No. Our first measurements were undertaken just to demonstrate the technique.

Q. (R. Hutchinson): What time resolution capability do you expect to attain?

A. (R. Murphy): The time resolution is limited by the time response of the detector. Our present capability is on the order of 10^{-6} sec.

Q. (D. A. Othen): Do you have any difficulty reproducing the relaxation phenomena at the different mirror positions used in your experiment?

A. (R. Murphy): The process of exciting CO_2 and CO has been used extensively by Legay et al. and in our own laboratory for a number of years. Our experiments show that the discharge and radiation from CO_2 under the conditions described is quite reproducible. Each interferogram point in Figure 31-3 is coadded over 10 discharges, which averages out any unwanted spurious signals.

Contents

32. Comments on Digital Data Acquisition Systems

Michael L. Forman
Goddard Space Flight Center
Greenbelt, Maryland

Abstract

The effect of digitizing noise is shown to be $E^2/3$ at each data point of an interferogram where $2E$ is the quantization voltage. For a fixed number of bits, a level slicing technique is demonstrated that reduces E by a factor of two or more in most cases.

A compact digital data acquisition system, suitable for field use and utilizing an analog tape recorder, is shown to have a signal-to-noise ratio far exceeding that of the analog recorder. The data rate is limited only by the high frequency cutoff of the direct record channels used.

General guidelines for designing data acquisition systems are discussed.

32-1 INTRODUCTION

Data acquisition systems are often the most neglected subsystems of an experiment and they are usually the last subsystems to be considered. Some experiments require very little in the way of data acquisition and a pencil or graphic recording device is all that is needed. Other experiments, which produce large volumes of data intended for computer processing or which require highly accurate data, usually have digital readout and recording devices associated with them.

The discussion in this chapter is concerned with the following topics:

(1) The noise introduced into the data by digitizing it, and a means of reducing this noise for a fixed number of binary bits.

(2) Some tangible and intangible aspects of a data acquisition system.

(3) A high speed data acquisition system that utilizes either digital or analog tape recorders.

32-2 DIGITIZING NOISE

Analog-to-digital converters are binary devices that quantize an analog signal into K levels, each of which is represented by a binary number. Let $K=2^k$, where k is the number of binary bits needed to represent a number, and let $2f(0)$ be the peak voltage. The latter is due to the fact that an analog function or interferogram $f(x)$ may be positive or negative with a maximum value of $|f(0)|$. Let $V(jE_a)$ be the assigned value for $f(x)$ if $jE_a<f(x)\leq(j+1)E_a$ when $j=0, 1, \ldots, K-1$, where $2E_a=2f(0)/K$. (See Figure 32-1(a).)

For a small sampling interval Δ, $f(x)$ may be approximated by a straight line of slope m as shown in Figure 32-1(b).[1] Thus the error due to quantizing is represented by

$$\epsilon(x)=mx \text{ where } -\frac{E_a}{m}<x\leq\frac{E_a}{m}. \qquad (32\text{-}1)$$

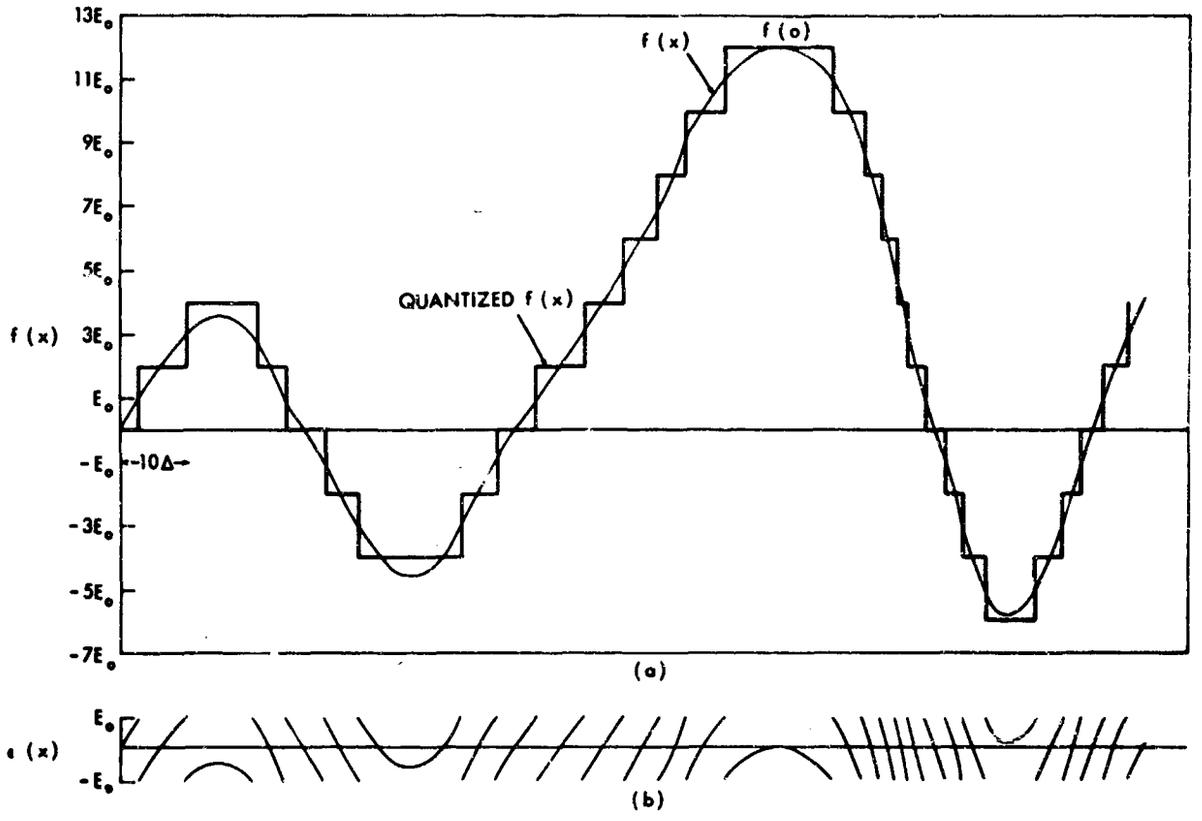


Figure 32-1(a). Staircase Quantizing; (b). Error Estimation Due to Quantizing

The mean square error is

$$\frac{m}{2E_o}\int_{-E_o/m}^{E_o/m}\epsilon^2(x)\,dx=\frac{E_o^{\,2}}{3} \qquad (32\text{-}2)$$

for each data point. The rms error voltage is then $f(0)/K\sqrt{3}$; hence, the signal-to-noise ratio in any analog signal for digitizing noise alone is

$$\left(\frac{S}{N}\right)_D=2^k\sqrt{3}. \qquad (32\text{-}3)$$

Another scheme for quantizing an analog signal or interferogram is to quantize points less than a certain percentage of full scale to 2^k and to use a coarser quantization for the remaining points. For instance, the Infrared Interferometer Spectrometer (IRIS) experiment aboard the Nimbus B satellite quantizes the points of the interferogram that are less than 10 percent of the maximum to 256 levels. The remaining points are then divided by 10 and quantized to the same number of levels, and the operation is recorded by one bit of a 10-bit telemetry data word. The mean square noise due to this is

$$y\frac{(E_o')^2}{3}+(1-y)\frac{E_o^{\,2}}{3}, \qquad (32\text{-}4)$$

where y is the fraction of total points in the interferogram that are digitized with $2E_o'$ as the step size. The coarse quantization step size is $2E_o$. Since

$$E_o=\frac{f(0)}{2^k} \text{ and } E_o'=\frac{pf(0)}{2^k}\quad p\leq 1,$$

the rms noise becomes

$$\frac{E_o}{\sqrt{3}} \cdot \sqrt{y(p^2-1)+1},$$

and the signal-to-noise ratio due to digitizing is

$$\frac{2^k \cdot \sqrt{3}}{\sqrt{y(p^2-1)+1}} = \left(\frac{S}{N}\right)'_D, \qquad (32\text{-}5)$$

where the denominator is referred to as the increase factor. Since y and p are always less than one, an increase in signal-to-noise ratio results with the same number of bits. For example, if 90 percent of the words are less than 10 percent of the maximum value, there is an increase of a factor of three, using the same number of bits that corresponds to a 2-bit gain. Figure 32-2 shows a plot of the increase factor as a function of p for several values of the fraction y. Such a scheme is useful when one has an 8-bit analog to digital converter but needs 10-bit accuracy. Figures 32-3(a) to 3(d) illustrate this effect on simulated data.[2,3]

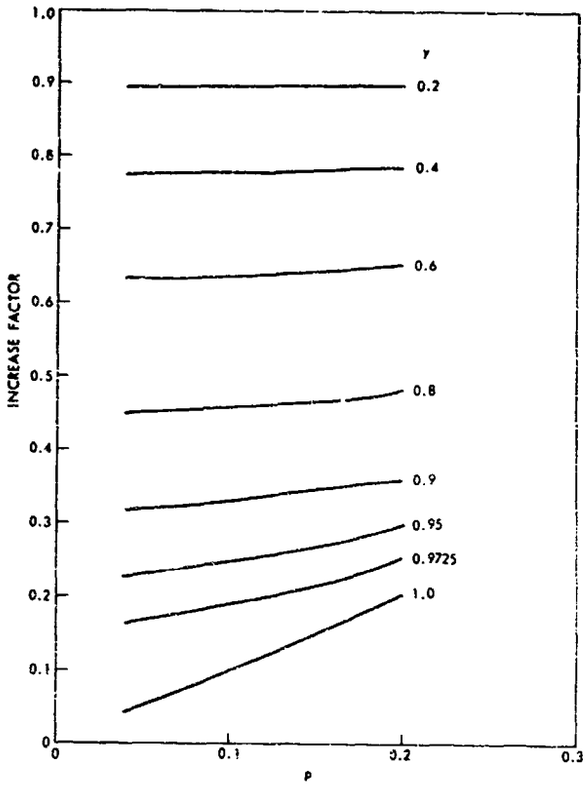


Figure 32-2. Signal-to-noise Increase Factor $\sqrt{y(p^2-1)+1}$ Vs p

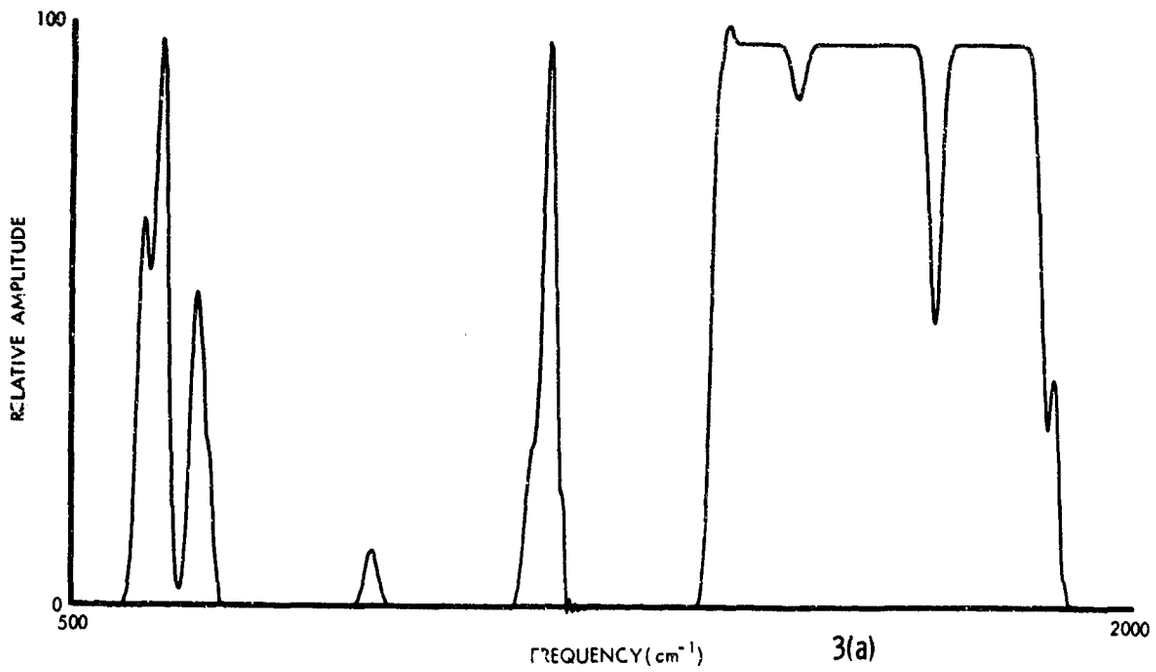


Figure 32-3(a). Effect on Simulated Data, No Quantizing

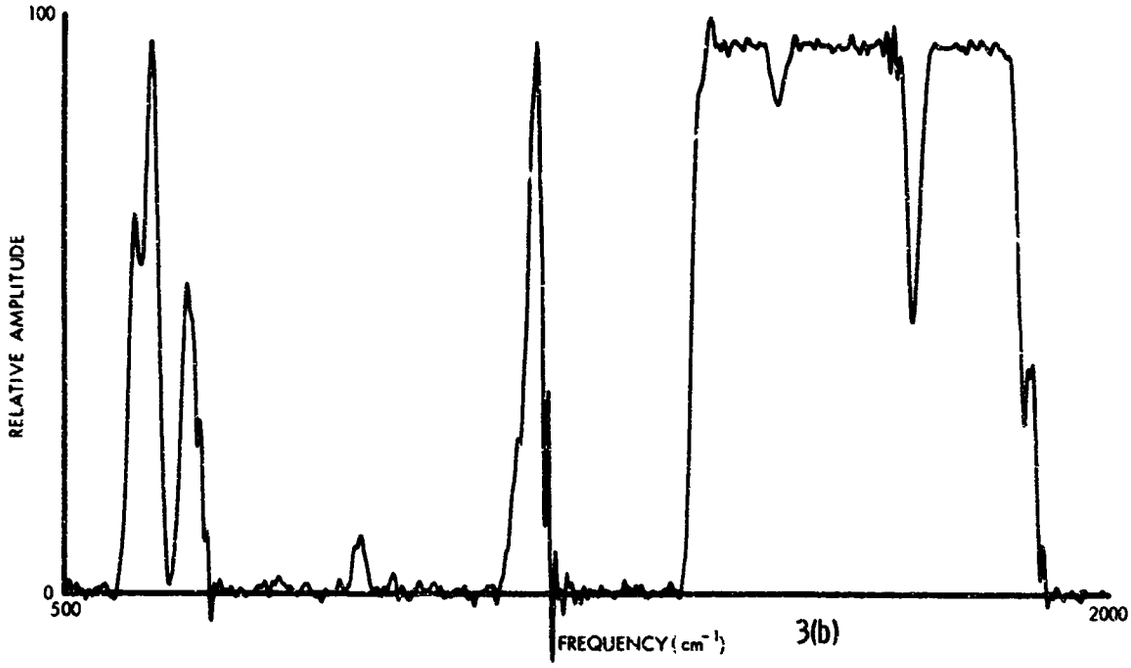


Figure 32-3(b). Effect on Simulated Data, Quantizing to 256 Levels

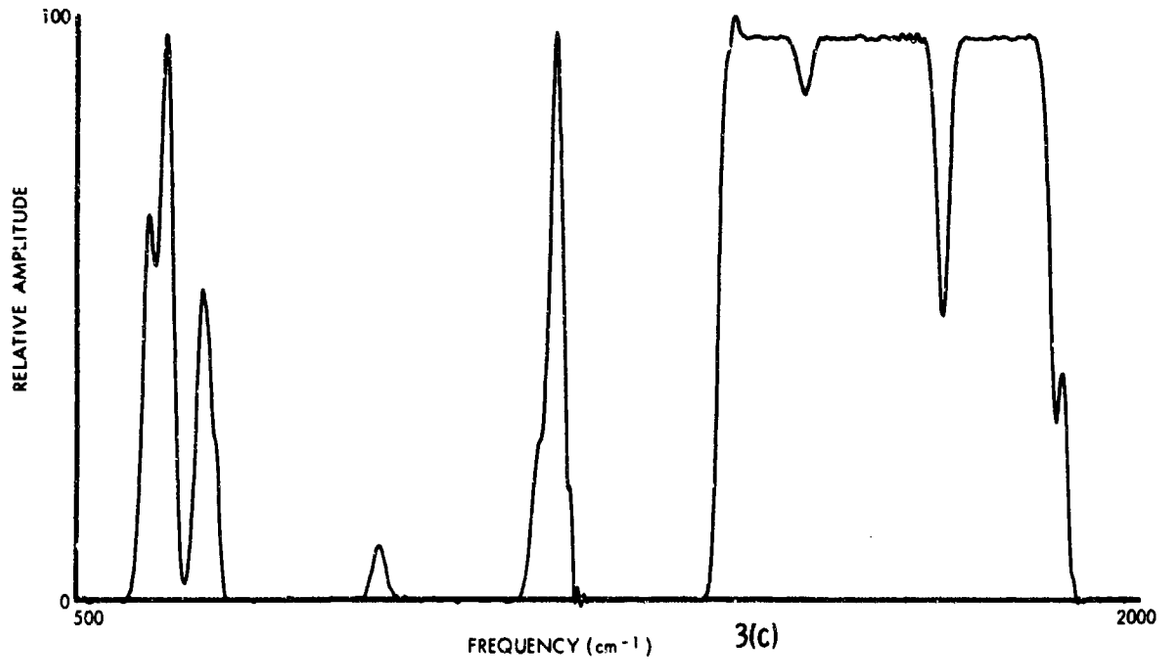


Figure 32-3(c). Effect on Simulated Data, Quantizing to 1024 Levels

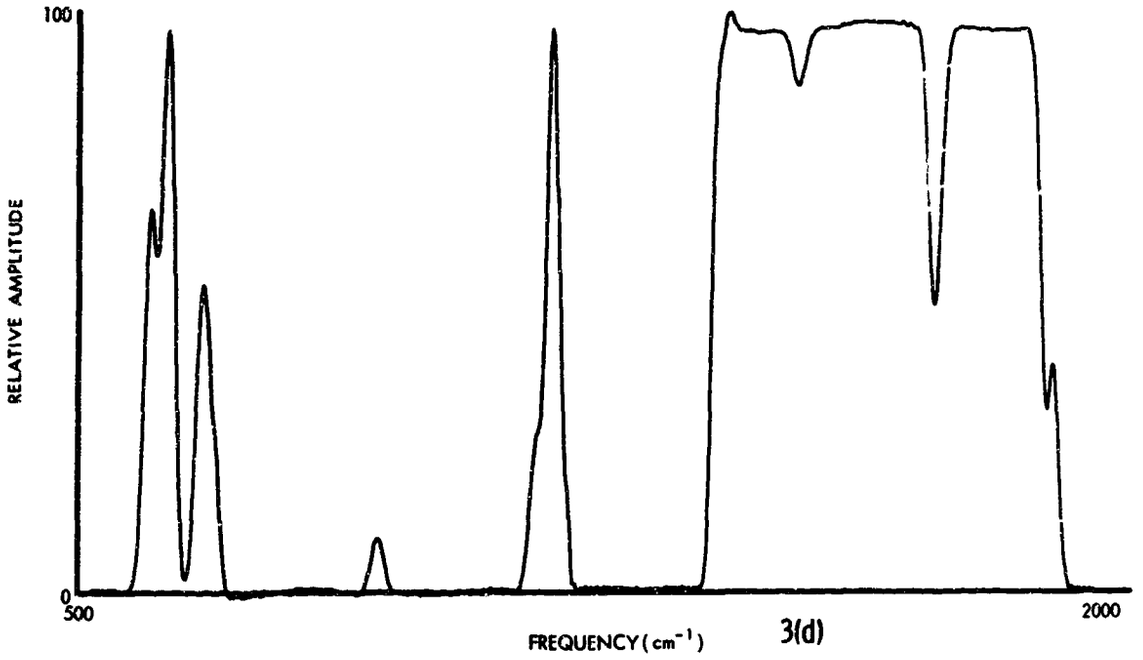


Figure 32-3(d). Effect on Simulated Data, Quantizing on a Divide-by-10 for 256 Levels. $0 \leq |f(x)| < .1f(o)$; $.1f(o) \leq |f(x)| \leq f(o)$

32-3 DIGITAL DATA ACQUISITION SYSTEMS—TYPES AND CONSIDERATIONS

32-3.1 Types

Digital data acquisition systems are usually either buffered or nonbuffered. The nonbuffered system is normally used for slow-data-rate applications and may consist of an analog-to-digital converter, an interfacing formatter, and a recording device such as an incremental tape unit (magnetic or paper). The buffered system may have the same components as the nonbuffered system with the addition of a data storage device such as a core memory. The primary advantages of the buffered system are the following:

(1) The throughput rate is much faster than that of the nonbuffered system because fast stop-start tape recorders may be used. The data are stripped from the memory at a rate consistent with the ultimate speed of the tape drive unit so that proper computer packing density is assured.

(2) Inter-record gaps are generated during acceleration and deceleration time of the tape recorder, and computer records are of uniform length.

(3) No data are lost, because the strip-out rate is much higher than the input data rate.

32-3.2 Design Considerations

There are certain system design considerations to think about before purchasing a system or its components. Some of them relate to component selection, and others to computer processing problems.

The first and perhaps the most overlooked item is the question of which computer is available for processing the data and whether it has inherent idiosyncracies due to its operating system. For example, with a system that produces signed binary data, one may have a computer that treats all binary data as signed two's complement. (See Table 32-1.) Of course if these facts are known, one can program around them. Other items such as record length, parity, number of tracks, tape density, and order of data being read, should also be considered in order to obtain a tape record format that is compatible with many computers within the framework of the available computer, and that also requires a minimum of manipulation.

The next items to consider are purpose and reliability. Will today's system work with next year's experiment with respect to data rate, computer to be

Table 32-1. Comparison of Signed Binary and Signed Twos Complement Representation (0 = +, 1 = −) for Four Bits Plus Sign

Decimal	Signed Binary	Signed Twos Complement
	S	S
+6	0 0 1 1 0	0 0 1 1 0
−9	1 1 0 0 1	1 0 1 1 1

used, and number of inputs and bits? It is far less expensive in both time and money to design now for the future. Where the system will be used, and whether environmental conditions and available power are amenable to the system, are also pertinent questions. With regard to reliability, everyone would like a zero defect system. Since this will not be available for some time, outputs should be monitored by digital-to-analog converters and compared to the original analog input. Small-scale computers can perform real time reduction and checks; however, finances, surroundings, and available personnel may make their use unfeasible. To help ensure continued good performance, specify reputable components and draw on the experience of those who are familiar with the type of system to be used. Some manufacturers of components tend to give overly optimistic performance specifications. To compensate for this, it is well to overdesign the system by a factor of two, especially with regard to data rate.

32-3.3 Tape Recorders and Analog-to-Digital Converters

A digital magnetic tape recorder is the link between the experiment and the computer. The quality and reliability needed in this component cannot be overemphasized. Its function is to present the computer with a readable tape every time it is used.

Incremental tape recorders are rated for a given packing density at a given number of characters per second. If a binary code is used to represent the data, a minimum of two characters are needed to represent the data word. If a tape record consists of fewer points than will be taken during an experimental run, the sampling rate without loss of data is governed by the end-of-record gap time (usually 25 to 80 msec) unless a compensating buffer is incorporated.

Fast start-stop continuous recorders, on the other hand, record data at a constant tape speed by stripping data from a memory at a rate consistent with the packing density required. The number of characters per second is equal to the packing density times speed. As previously mentioned, end-of-record gaps are generated during the start-stop times (about 5 msec) and proper utilization of the data buffer avoids any data loss. Both types of recorders can be used with asynchronous sampling rates.

Analog-to-digital converters (A/D) and multiplexers make it possible to digitize one or several channels of information either sequentially or simultaneously, depending on the number of sample hold amplifiers used. The A/D converters are usually rated in time per bit; however, this factor may be degraded when used in conjunction with a multiplexer. Thus, when specifying the data acquisition rate, either in words or bits per second, do it for the total system.

The number of bits per data word required is a function of the digitizing signal-to-noise-desired. A general guideline is to digitize to two bits more than the experiment's expected signal-to-noise. The accuracy of A/D converters is expressed as a percentage of full scale (0.01 percent) plus or minus one half of the least significant bit.

32-4 A VERSATILE SYSTEM

This section describes a general-purpose high-speed data-acquisition system built for the Goddard Space Flight Center by Digital Products Corporation.

The system consists of two subsystems: the first acquires the data at prescribed rates, and the second converts this data to computer-compatible form. These two subsystems may be used together or with an analog tape recorder between them. Six numeric thumbwheel switches and the output of a time code generator provide identification.

The system accepts up to eight inputs that are programmable at a sequential sample rate of up to 92 160 ten-bit words per second. All available sampling rates (binary dividers of the maximum rate) are controlled by an internal clock. The bit stream produced by the A/D may be recorded directly on a computer-compatible tape if the output bit rate is less than 122 kHz. Alternatively, at any data rate, the bit stream can be recorded on one or more direct record tracks of an analog tape recorder. If the frequency response of a track is exceeded, the output bit stream is multiplexed on a frame basis so that the bit rate to each track is below the high frequency cutoff of the analog track. A frame of data is illustrated in Figure 32-4. Frame multiplexing is chosen in order to

SYNC BITS	HEADER TIME	120 TEN BIT DATA WORDS
1-20	21-80	81-1280

Figure 32-4. Data Frame Layout

assure adequate identification and to eliminate tape skew problems. The analog tape may then be played back into the unit, which produces the computer-compatible tape.

If recording is done on analog tape, a split phase binary code (Figure 32-5) is used rather than a code where X volts represents a one, and Y volts represents a zero. This is done because the low frequency re-

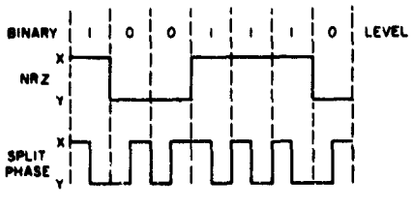


Figure 32-5. Bilevel Coding (upper) vs Split Phase (lower)

sponse of direct record channels is poor, and a string of ones and zeroes will cause drift problems. The split phase format assures one transition per bit regardless of the sequence of ones and zeroes.

Calibrated voltage sources are incorporated into the system to check the performance of all channels. Digital-to-analog converters are available at several points within the system to provide visual checks of the digitized data. Figure 32-6 is a block diagram of the data acquisition system.

This system is to be used to gather data from a variety of infrared sensors, some of which require extremely high sampling rates. It is adaptable to field use in confined areas such as small aircraft, or to use in the laboratory. After a minimum of setting switches, the user is free to concentrate on important items. Whether or not an analog tape recorder is used, the signal to noise is limited by the digitizing system, not by the recorder.

32-5 APPLICATION TO FOURIER SPECTROSCOPY

With a minimum of change, the system will accept asynchronous sampling rates such as those produced in Fourier spectroscopy. For this purpose the A/D should be changed to give higher resolution. Without change, the system can be used applying the technique of Sakai and Murphy[4] in which both the interferogram and fringe reference are digitized at dense sampling rates on a constant time basis.

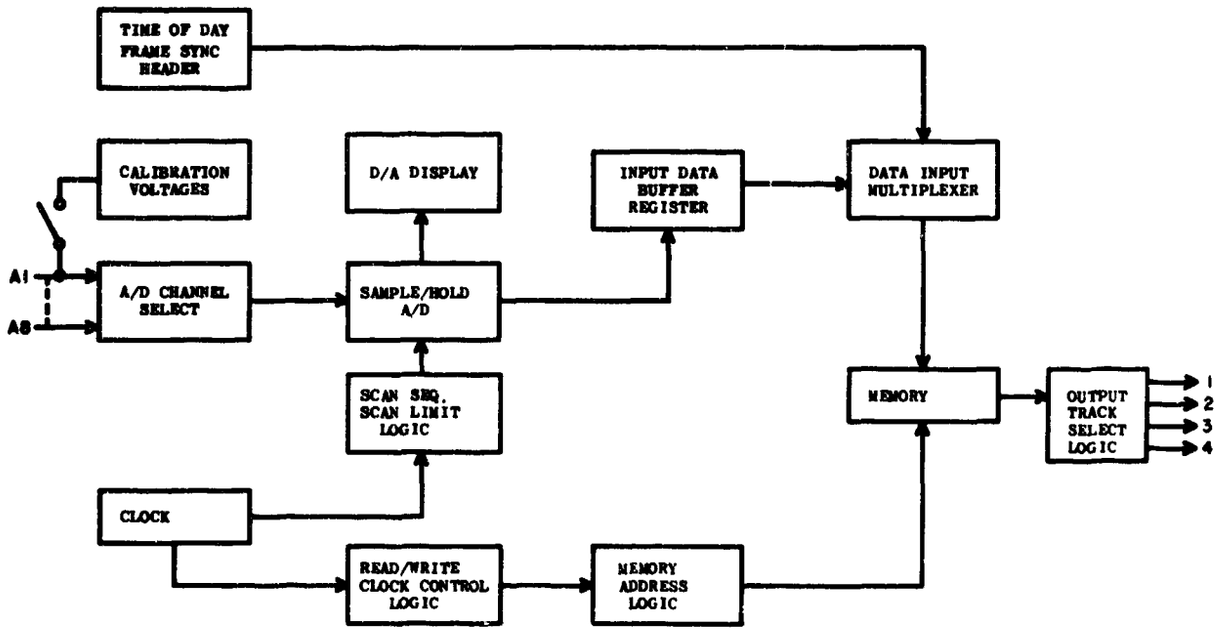


Figure 32-6. Block Diagram of Data Acquisition System. (Ground formatting system is not shown)

References

1. Bennett, W. (1948) *Bell Sys. Tech. J.* **27**: 446.

2. Steel, W., and Forman, M. (1966) *J. Opt. Soc. Am.* **56**: 982.

3. Forman, M. (1966) NASA-GSFC Document X-622-66-478, Sept. 1966.

4. Sakai, H., and Murphy, R. *J. Opt. Soc. Am.*, to be published.

Discussion

Q. Do you see the surface of Venus at these wavelengths?

A. (R. A. Hanel, Goddard Space Flight Center): No.

Contents

33. Real-Time Computer for Fourier Spectroscopy

P. Connes and G. Michel
Laboratoire Aimé Cotton, C.N.R.S.
Orsay 92 France

Abstract

A real-time classical Fourier transform computer for Fourier spectroscopy with a basic computation time of 10^{-6} sec per input point per output point has been built. When used with the Aimé Cotton Laboratory interferometers it gives 20 000 spectral samples from interferograms recorded at 50 samples per sec. It is easily adaptable to high or low resolving power and broad or narrow spectral range. No large random access memory is needed and the instrument is simpler and cheaper than the smallest general purpose computers, which are approximately 10^3 times slower for this particular purpose.

33–1 INTRODUCTION

In the last few years an amazing number of Fourier analyzers or computers for the purpose of analyzing time functions have been described or put on the market. On the other hand, programing of general purpose computers has progressed with such giant strides that one can well ask whether Fourier spectroscopists should continue their efforts to produce their own analyzers. Before answering, one must study their needs and all the available devices.

Chapters 6 and 11 by J. Connes and H. Delouis summarize the situation as far as *a posteriori* analysis with large computers goes. They solve the problem of transforming the largest number of samples of practical interest (for example, 10^6 samples in 9 min). It is obvious that unspecialized commercially available computers, which are small and cheap enough to be considered as auxiliaries to the interferometer, can solve the problem in the far more common case where $N \leq 10^4$; medium size computers should be able to handle the intermediate case where $N \leq 10^5$. In this paper N will designate the minimum number of samples needed for a perfectly odd or even interferogram recorded from $\delta = 0$ to δ_{max}; $M \leq N$ will be the number of spectral samples wanted, with wavenumber separation $\delta\sigma = 1.2\delta_{max}$.

Real-time computation is a different problem. While of course somewhat of a luxury, it is nevertheless extremely useful when the recording time is very long or very valuable, as in astronomical observations for instance. The need for real-time computation also increases with the complexity of the interferometer system since the probability of wasting valuable recording and computing time increases. This is particularly true of the high-resolution, large-spectral-range interferometers set up at Aimé Cotton Laboratory.[1] (See Chapter 12 by Guelachvili and Maillard.) In this case, however, the information content of the spectrum is so large (up to 10^6 samples) that it becomes unnecessary to compute the entire range in real time; if available it could not possibly be examined in real

time (except by multiplexing the experimenter). A well selected spectral slice containing M spectral elements provides all the elements needed for checking resolution, signal-to-noise (S/N) ratio, systematic errors, and so forth. A spectral slice of $M > 10^3$, that is, not too narrow, should be used. Operation should be compatible with a recording speed of 50 samples per sec, an interferogram dynamic range of the order of 10^6, and a spectral dynamic range of the order of 10^4. Preferably the final (full resolution) spectrum should be apodized and accurate interpolation between spectral elements should be performed. No available device comes anywhere near meeting such requirements.

33-2 DISCUSSION OF REAL-TIME COMPUTERS AND ANALYZERS*

Devices will be classified according to whether the stored quantity is the time function (in the case discussed here, the interferogram) or the spectrum. They may be analog, digital, or hybrid, and may differ widely in speed and accuracy; of course they all obey the basic uncertainty relationship, $F = 1/T$, where T is the recording time and F is the frequency resolution. For the case discussed here, the most important consideration is the computation time as a function of the number of output samples; a general comparison is shown in Figure 33-1, which gives the computation time vs the number of spectral samples M for the following analyzers or computers:

(a) FFT Computers. The number of input samples N is equal to M; computation is a posteriori. The representative curve differs little from a unit slope straight line. The IBM 360/75 is a general purpose computer programed by Delouis (see Chapter 11). The two straight lines correspond to operation with or without disks. The IBM Real-Time Digital Analyzer (RTDA) is a special purpose computer attachment implementing the FFT in hardware (must be connected to an 1800 or 360 computer). The Hewlett Packard 5450 A must be used with general purpose computer 2115 A. The Computer Signal Systems CSS-3 is a completely autonomous FFT system. The Time Data 100 (Time Data Corp.) is a completely autonomous FFT system giving 1000 samples in 1 sec.

(b) Time Compression Systems (slope 2 straight line). Federal Scientific Corp. ubiquitous analyzer and Signal Analysis Industries Corp. real-time analyzer are both approximately represented by the same line. The Research Industrial Instruments Company time compression analyzer was built especially for Fourier spectroscopy.[12]

(c) Electrooptic Spectrograph.[5]

(d) Sinoramic Analyzer (Federal Scientific Corp.). Each line corresponds to a different instrument.

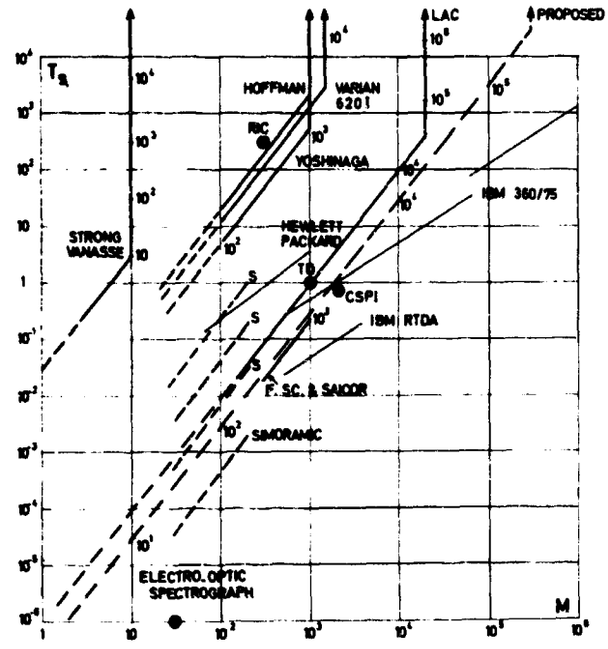


Figure 33-1. Computation Time Vs Number of Spectral Samples M for Different Analyzers or Computers

(e) Real-time computers that use the classical FT algorithm: Operation along oblique straight line with slope 2 means $N = M$ and the recording rate is limited by the computer speed. Operation along vertical line means N and T can increase without limit, M is limited by computer capacity, N is indicated by numbers to the right of the lines. Strong-Vanasse: Electromechanical computer.[10] Varian 620 I: General purpose small computer, as programed by F. Levy et al (Chapter 34), Hoffman,[2-4] and Yoshinaga:[11] specialized real-time Fourier spectroscopy computers. Laboratoire Aimé Cotton: computer described in this chapter. A possible later version with a speed increase of a factor of 3 is proposed. This could be built using presently available components. Storage of the spectrum would use disk or drum memory.

There should also be mentioned the association of a hybrid real-time autocorrelator and of an analog Fourier analyzer (Princeton Applied Research) which would provide the square of the optical spectrum. The correlator produces 100 samples of the autocorrelation function of the input; the multiplications are performed by hybrid techniques and the 100 results are stored in analog form as capacitor voltages. Maximum delay time is 10 sec. The fully analog Fourier analyzer scans the 100 capacitor bank once for measuring each spectral element.

33-2.1 Digital Computers Using the FFT Method

Digital computers using the FFT method should be called "false" real-time analyzers; the FFT algorithm cannot be applied until all the data samples have been collected. Computation time is totally

* This review is limited to devices which either have present-day interest or are closely linked to the history of Fourier spectroscopy. A more complete survey is given by Hoffman.[2]

independant of recording time. However, some of the special purpose computers with a wired-in FFT algorithm are so fast that they might be used for Fourier spectroscopy by: (1) separating the N input samples into K blocks of N/K samples each; (b) computing a first approximation spectrum from the first N/K samples after recording of the first block is complete; and (c) computing a new spectrum from the $2N/K$ available data after the second block, and so forth. The check on interferometer performance would be discontinuous, which is no great disadvantage. A similar technique has been applied[5] to a time-shared IBM 1800 computer. The maximum number of input samples treated appears to be of the order of 10^3. No indication of computing speed is given.

The most severe limitation comes from memory size; a random access memory is essential. The number of outputs M is automatically equal to the number of inputs N and this is 4096 in the highest capacity special purpose FFT computer built so far.

Thus, direct checking of a high N interferogram is impossible, but the difficulty is not fundamental. One could proceed in two stages with a real-time program similar to the Test 3 program described in Chapters 6 and 11 and used so far for a posteriori rapid checking of interferograms. The recorded data would be convolved by a stored "filter" function producing a "secondary" interferogram with only $n \ll N$ samples; n should be small enough to fit in the memory, which would slowly fill up as recording proceeded. The primary interferogram does not have to be stored. The normal FFT program, applied at intervals, would show the wanted spectral slice at progressively increasing resolution. Both operations can be performed in either general purpose or specialized computers; however, the overall program is complex and might prove difficult to handle with the smaller machines. Lastly, if the available computer has external nonrandom access memories (tape or disk) a real-time program derived from the FFT 1000 K could also be used (see Chapters 6 and 11).

33-2.2 Hybrid Analyzers Using Time Compression

In hybrid analyzers using time compression, just as for FFT computers, the stored function is the interferogram; however, unlike FFT computers, they are true real-time analyzers.

The interferogram is first digitized, and then stored in a circulating memory (MOS transistors shift register or quartz ultrasonic delay line). It is then extracted, sent through a digital-to-analog converter, and frequency-scanned by an ordinary wave analyzer. Analysis is repetitive and can, but does not have to, proceed simultaneously with recording; one spectral element per recirculation is measured.

The number N' of input samples is limited by memory size. Since analysis is not phase sensitive, a power transform is obtained. This means the number of input samples N' should be at least $2N$; in fact $N' = 3N$ is chosen, to provide some oversampling.

The time needed to scan the full memory is equal to the recirculation time and bears no relation to the recording time. It can be made much smaller, that is, time compression is used. The speed limitation lies in the minimum memory readout time, which is of the order of 10^{-7} sec per word (about 8 bits recorded in parallel). If M is the number of spectral elements, the analysis time is $T = MN' \times 10^{-7}$ sec. Since this is a general relationship which applies to all systems not using the FFT, using the minimum number of input samples N:

$$T = MN\theta$$

with

$$\theta = 3 \times 10^{-7} \text{ sec.} \qquad (33\text{-}1)$$

In Figure 33-1 this class of devices is represented by a slope 2 straight line. Maximum capacity so far is $N' = 3 \times 10^3$ and $M = 10^3$. To process the high N interferograms, word capacity should be increased by a large factor, "proposed" in Figure 33-2. Use of a drum

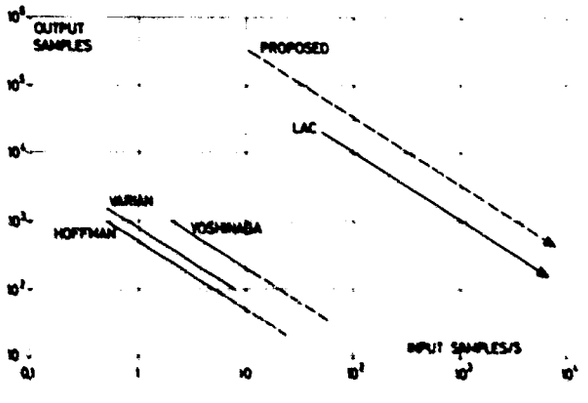


Figure 33-2. Number of Output Samples That Can Be Computed at a Given Input Sample Rate by Different Real-Time Computers

or disk store, plus a buffer, is a conceivable extension; the basic $T = MN\theta$ law would still apply. The number of bits per word should also be increased, however, and the accuracy of the analog components would then prove inadequate.

33-2.3 Analog Electrooptic Spectrograph

The analog electrooptic spectrograph, an interesting device that was built for radio astronomy and that

is obviously unsuitable for present day Fourier spectroscopy, will only be mentioned. It is a true real-time analyzer in which the input time function is stored as a traveling wave in a transparent medium, thus producing a phase grating. The diffraction pattern of a monochromatic line parallel beam gives a power spectrum. Ultimate limitations are not clear; 30 spectral elements have been analyzed in about 10^{-6} sec (30 kHz resolution in a 1 MHz total bandwidth).

33-2.4 Analog Analyzer With Spectrum Storage (SIMORAMIC)

The analog analyzer with spectrum storage is a new class of device in which the spectrum itself, not the interferogram, is the stored quantity. Thus there is no limitation on the number of input samples, a very important point for the application discussed here. Whether analog or digital, these devices are perhaps best understood by using the language of sampled functions and remembering the classical FT algorithm. As each new input sample becomes available it is multiplied by M suitably chosen samples of the sine function and the results are put into M storage locations. The next input sample is treated in the same manner (with different sine samples), the old results are extracted from the store, added to the new, displayed on an oscilloscope, and stored again. Looking ahead to Figure 33-19, one can see how resolution increases in this manner.

The SIMORAMIC[7] analyzer is fully analog, and no sampling is actually used. The store is an analog recirculating delay line. Sine generation is performed by a local oscillator and multiplication by a frequency changer. Just as in the time compression systems, a full spectrum is displayed for each recirculation.

Speed is limited by bandwidth of line and transducers, and the (equivalent) number of output samples by line capacity, that is, length. Several instruments with the same capacity ($M = 200$) and recording-analyzing time in the range 2×10^{-3} sec to 1 sec have been built. They are intended to be used at full output capacity; however, it is interesting (for comparison purposes) to note that operation of a given instrument at reduced line length would give a proportional reduction in analysis time T and in the number of samples $M \leq M_{max}$. Since recording is simultaneous, the number of samples needed to describe the input signal would be reduced in the same way. For the fastest model $T = 2\times10^{-3}$ sec at $M_{max} = N_{max} = 200$. Thus

$$T = MN\theta$$

with

$$\theta = 5\times10^{-8} \text{ sec.} \qquad (33\text{-}2)$$

However, unlike case 2, distortions set a practical limit to the number of recirculations, and thus to storage time ($T \simeq 1$ sec). Unless preliminary time compression were used, longer duration interferograms could not be handled. Fellgett[8] considered possible applications to Fourier spectroscopy and showed that any improvement would inevitably lead to digitalization, a conclusion that is fully borne out by the present paper. Manufacture of the SIMORAMIC analyzer appears to have been discontinued.

33-2.5 Mechanical or Electromechanical Computers

The venerable ancestor of mechanical or electromechanical computers is of course Michelson's *"New Harmonic Analyser,"* shown in Figure 33-3. Up to 80 sine waves of adjustable amplitudes and phases could be added to synthesize any function, of which many examples were given. (See Figure 33-4.) Operation for analyzing a curve was not automatic and required manual adjustments in each channel; thus it would be hard to quote a figure for speed of operation. Michelson was obviously pleased with the results ("The experience gained in the construction of the

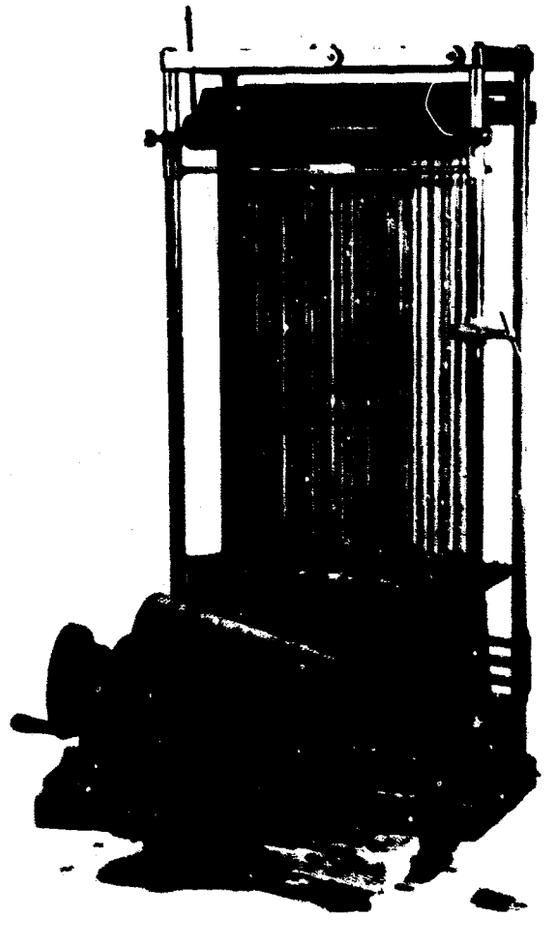


Figure 33-3. All Mechanical Fourier Analyzer Built by Michelson (Michelson and Stratton[9]). Photograph by courtesy of the Smithsonian Institution, Washington, D.C.

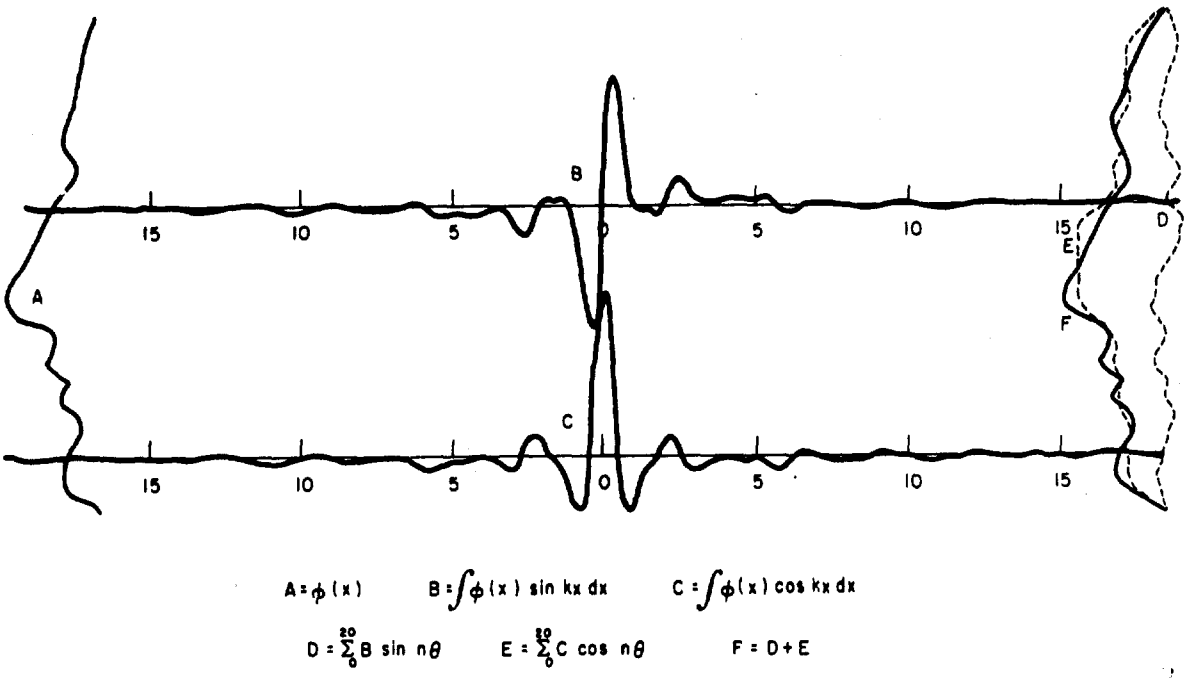


Figure 33-4. Results Given by Michelson's Analyzer. Left: Artificial given "spectrum" (vertical frequency scale). Center: sine and cosine "interferograms" computed by the analyzer itself, and limited to 20 samples. Right: sine and cosine transforms of the two "interferograms" their sum approximately reproduces the original "spectrum"

present machine shows that it would be quite feasible to increase the number of elements to several hundred or even to a thousand with a proportional increase in the accuracy of the integrations"), but does not seem to have actually used it for analyzing visibility curves.

The first real-time analyzer for Fourier spectroscopy in the modern sense was described by Strong and Vanasse.[10] The input is not sampled but "multiplied" by M switches, the "sine functions" being generated by gears linked to the interferometer motion. The spectrum is sampled, since M discrete outputs are given, by M mechanical (ball and disk) integrators. Since these are drift-free, there is at last an instrument with the capability of handling interferograms of infinite duration and length in real time, and which is suited to checking a spectrum fraction at full resolution.

Cost and size limit the number of channels to $M_{max} = 10$ and the instrument was intended for use at full capacity, that is, $N > M_{max}$. Analyzing-recording time grows with N (vertical line in Figure 33-1). However, for comparison, consider a situation where the number of output points should be reduced due to finite computer speed; then $M = N$. Since, in the instrument as described the bandwidth for each channel is 0.4 Hz, the minimum recording time is 2.5 sec for $N = M_{max} = 10$. Thus, for $N = M < M_{max}$

$$T = MN\theta$$

with

$$\theta = 2{,}5.10^{-2} \text{ sec.} \qquad (33\text{-}3)$$

Operation in this region is represented by a slope 2 straight line. Note that this representation applies to an instrument of given output capacity. Adding more parallel channels but keeping the same speed limitations would translate the slope 2 straight lines horizontally to the right on Figure 33-1. The same basic limitations will be found in the much faster electronic computers to be described.

33-2.6 Hybrid Real-Time Computers for Fourier Spectroscopy

Hybrid real time computers for Fourier Spectroscopy embody in a direct manner the classical FT algorithm according to the general scheme described in Section 33-3.4.

33-2.6.1 THE YOSHINAGA ET AL COMPUTER[11]

The Yoshinaga et al computer is a hybrid, predominantly digital, type. For apodization the interferogram is first weighted by an analog programer, and then sampled and digitized. The sine function is produced by a hybrid generator; multiplication and

addition are performed by an incremental digital technique. Storage of the spectral samples takes place in a random access core memory; this is needed because of the particular addressing scheme used. Thus, capacity has been limited to $M_{max}=1000$. The time taken to compute the full 1000 outputs from 1 input is 0.5 sec. in the $N=M<M_{max}$ case, therefore

$$T=MN\theta$$

with

$$\theta=5.10^{-4}\text{ sec.} \qquad (33\text{-}4)$$

Accuracy should be good, although it is somewhat difficult to assess from published results. The instrument was built for far-infrared spectroscopy and it is used to compute the full spectrum. A modified version could compute a spectrum slice from a high interferogram, but speed would become a severe limitation. Only 40 output samples could be computed at 50 input samples per second.

33-2.6.2 THE HOFFMAN-VANASSE COMPUTER[2,3,4]

The Hoffman-Vanasse computer, also built for far-infrared work, is another example of hybrid, predominantly analog, techniques. The interferogram is weighted for apodization, and then sampled, but in analog form. Sine generation is performed by a frequency synthesizer and multiplication is analog. Only then is the result digitized. Addition and storage are purely digital (giving unlimited recording times) and digital-to-analog conversion produces an oscilloscope output. All these last operations are done by a commercial core memory signal averager.

Intensity accuracy is about 2 percent, mostly because of the analog components. Accurate timing of the sine function maxima is needed, and this is also a delicate point. Speed is about four times less than that of the more complex Yoshinaga computer:

$$T=MN\theta$$

with

$$\theta=2\times10^{-3}\text{ sec.} \qquad (33\text{-}5)$$

33-2.7 General Purpose Digital Computer

The classical FT algorithm for real time computation is basically the same as the one commonly used for a posteriori computation except that the input samples, not the output samples, are treated in sequence. But computers small and cheap enough to make this practical have appeared only lately. F. Levy et al in Chapter 34 describe how, with a Varian 620 I of only 4096 memory words, they have succeeded in computing $M_{max}=1500$ samples in 2 sec. Thus:

$$T=MN\theta$$

with

$$\theta=1.3\times10^{-3}\text{ sec.} \qquad (33\text{-}6)$$

This result is obviously an important one and the technique is adequate for many Fourier spectroscopists. In the far infrared the entire spectral range can be computed this way. However, the system is still much too slow for the purpose discussed here. No order of magnitude gains are in view, but a two-stage program involving the computation of a filtered secondary interferogram by convolution (Section 33-2.1) gives a sizeable improvement. Output capacity remains limited by the expensive random access memory.

33-3 SPECIAL PURPOSE DIGITAL COMPUTER

33-3.1 General Description

The special purpose digital computer involves no radically new principle. It is simply a digital computer with a wired-in classical FT algorithm. Analog devices are used only at the output for interpolation between spectral samples (and also apodization) in order to present a continuous curve on an oscilloscope.

The relative simplicity, high speed ($\theta=10^{-6}$ sec), and high capacity ($M_{max}=20\,000$ in the final, not yet completed version) are derived from three features:

(a) elimination of all programming,

(b) putting the accuracy where it is really needed. In practice this means the sine function can be stored in a relatively small core memory, with a very simple addressing scheme. It is common practice to transform odd interferograms recorded by the internal modulation technique.[1] (See Chapter 12.) The first sample can easily be taken at $\delta=0$ with great accuracy. Even interferograms can be transformed with a minor modification of the addressing scheme. Power transforms could be handled at the cost of a factor of 2 reduction in speed and capacity since two multiplications are needed, and two products should be stored. Squaring and adding for presentation of the results would be done most easily by analog means.

(c) Storing the spectrum in a sequential access circulating memory which is much cheaper per bit than a random access one.

Figure 33-5 gives the basic computer block diagram and the interconnections to the interferometer, which is described elsewhere.[1] (See Chapter 12.) The interferometer system includes its own digitizer, the output of which is sent in BCD form to the tape recorder for a posteriori computation by a general purpose computer (see Chapter 11). A second output, in binary form (14 bits), is sent to the real-time computer.

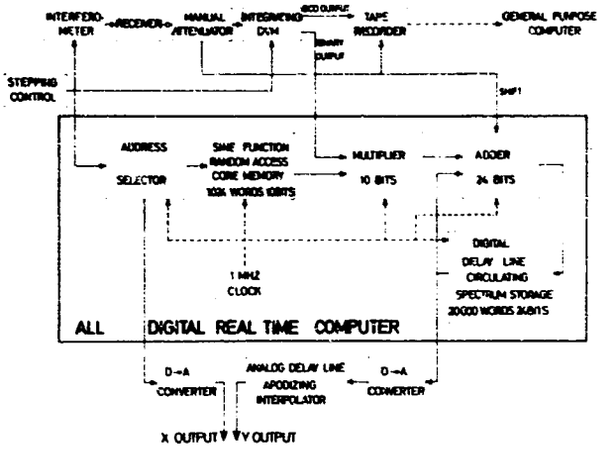


Figure 33-5. General Diagram of Real-Time Computer With Input and Output Connections

The interferometer stepping control delivers one pulse I after each interferogram sample has been recorded. Each pulse I triggers the operation of an address selector (Section 33-3.3), which finds in sequence the M addresses for the M needed sine function samples corresponding to the particular I_n interferogram sample received and to the M wanted spectral points. Operation is timed by a 1-MHz internal clock; thus 20 000 addresses are found in 20 msec and computation can proceed at a maximum rate of 50 input samples per second. The computer is totally asynchronous and would operate equally well with a stepping interferometer.

The sine function samples are extracted from a 1024 words, 10 bits, read-only, random access core memory with 0.6 μsec cycle time. Memory size bears no relation to the number of input or output samples. It limits only the accuracy of computed intensities (see Section 33-3.2).

The product of the interferogram and sine samples are computed by a 10 bits digital multiplier clocked at the same rate (10^6 multiplications per second). Handling of the extra bits to give 24 bits dynamic range is explained in Section 33-3.4. The 10 bit output is sent through a 24 bit adder to a 24 bit, 20 000 word delay line store made of 24 parallel 20-msec magnetostrictive delay lines; there are 50 recirculations per second. Input and output are again clocked at 10^6 words per second.

The outgoing words are sent back to the adder input. They also give the final computer output. A digital-to-analog converter provides an oscilloscope Y output, through an analog delay line apodizing interpolator. See Section 33-5.2. The address computer gives the X output through a second digital-to-analog converter. Thus the M samples spectrum can be presented at a 50-Hz repetition rate. A slowed-down pen recorder output is also available.

33-3.2 Quantization of the Sine Function

In the most general case MN samples of the sine function are needed for a N input M output transform. Considerable simplification results from the use of a simple addressing scheme. (See Section 33-3.3.) Only one quarter of a sine function period needs to be stored in sampled and quantified form. This, however, implies an approximation, and it is necessary to evaluate the rms error involved. The different addressing scheme of the Yoshinaga computer[11] does not basically involve any approximation; however, the actual accuracy in the multiplications is the same (10 bits). It is not applicable to the problem discussed here, for several reasons. A more expensive random access memory is needed for spectrum storage. The spectral and interferogram sampling intervals are linked by a rigid relationship, $\sigma_1 \delta_1 = 1/2503$. Even with suitable modifications the selection of any spectral range and spectral sampling interval would be difficult. Speed appears limited chiefly by the hybrid sine function generator.

Consider the desired function $y = \sin(\pi/2)x$ with $0 < x < 1$ (Figure 33-6). We can store only a finite

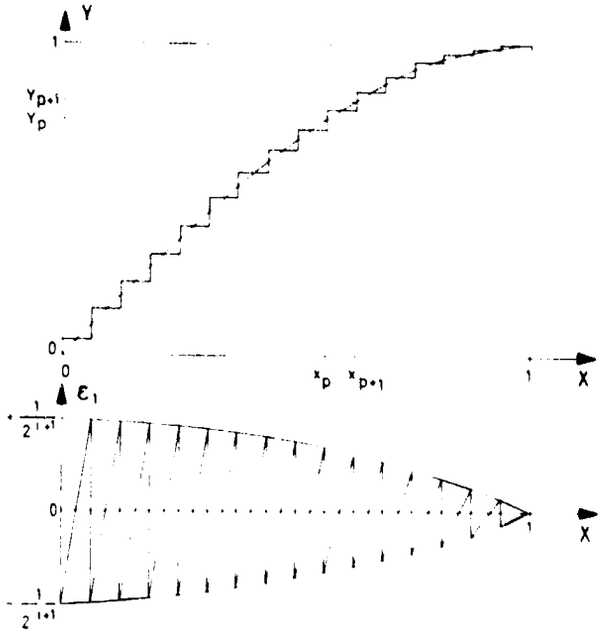


Figure 33-6. Quantization of Sine Function

number 2^i of y samples, each with a finite number of bits k. Thus, in fact two approximations are made, which introduce two different errors ϵ_1, ϵ_2. If the quantification is fine enough it is reasonable to treat the two errors as statistically independent, and the rms resultant error $\bar{\epsilon}$ will be

$$\bar{\epsilon} = \sqrt{\overline{\epsilon_1^2} + \overline{\epsilon_2^2}}. \tag{33-7}$$

33-3.2.1 FIRST ERROR

In the interval $x_p < x < x_{p+1}$, replace y by $(y_p + y_{p+1})/2$. The error is

$$\epsilon_1 = y - (y_p + y_{p+1})/2 \simeq \cos x_p \left[(x - x_p) - \frac{1}{2^{i+1}} \right], \tag{33-8}$$

where $p = 1, 2, \ldots p_{max} = 2^i$.
The error is illustrated by Figure 33-6. Finally,

$$\bar{\epsilon}_1 = 1/2\pi \cdot 2^i. \tag{33-9}$$

33-3.2.2 SECOND ERROR

Replace the y_p by their quantified values z_p. Use k bits, that is, 2^k binary levels. The error $\epsilon_2 = z_p - y_p$ is a saw-tooth function of constant amplitude $1/2^{k+1}$ and

$$\epsilon_2 = 1/4 \cdot 2^k. \tag{33-10}$$

For i and k of the order of 10, halving ϵ_1 means doubling the memory word capacity while halving ϵ_2 means adding one more bit, thus increasing bit capacity by 1/10 only. It is logical to take $\epsilon_1 > \epsilon_2$; however, memory price is not a simple function of word or bit capacity and an optimum solution is hard to define. Table 33-1 gives the rms error for a "reasonable" set of solutions.

The main limitation is the fact that the error decreases only linearly with word capacity and very high accuracy would become too expensive. A large improvement in accuracy for a given memory size could be realized by more complex addressing schemes. For instance, the memory can be divided into two

Table 33-1. Rms Error For a "Reasonable" Set of Solutions

Number of Words 2^i	"Optimum" Number of Bits k	Rms Error
256	9	8.10^{-4}
512	10	4.10^{-4}
1024	11	2.10^{-4}
2048	12	1.10^{-4}
4093	13	$0.5\,10^{-4}$

sections, the first holding ordinates and the second slopes. The sine function is approximated by secants and not horizontal lines. But the range of precision needed for Fourier spectroscopy can be realized with a moderate capacity. The first results (Figures 33-17 to 33-21) have been recorded with a temporary 256 word, 8 bit memory. The final version will include a 1024 word, 10 bit memory (11 bits not being available), that is, $i = 10$ and $k = 10$.

The distortion in the computed spectrum can be understood in several ways. For instance, in the case of a single monochromatic line σ_0 giving an interferogram $I(\delta) = \sin 2\pi\sigma_0\delta$, the normal integral

$$B(\sigma) = \int I(\delta) \sin 2\pi\sigma\delta \, d\delta$$

will be replaced by

$$B'(\sigma) = \int I(\delta)[\sin 2\pi\sigma\delta + \epsilon(\delta)] \, d\delta = B(\sigma) + \int \sin 2\pi\sigma_0\delta \cdot \epsilon(\delta) \, d\delta. \tag{33-11}$$

Since $\epsilon(\delta)$ is a periodic function of δ, it is possible to consider the terms of the equivalent Fourier series, the frequencies of which will be $\sigma, 2\sigma, 3\sigma, \ldots$. Thus, the integral $\int \sin 2\pi\sigma_0\delta \cdot \epsilon(\delta) \, d\delta$, which represents the error term in $B'(\sigma)$ will give peaks for $\sigma = \sigma_0/2$, $\sigma = \sigma_0/3$, and so forth.

Subharmonics of the line σ_0 will appear in the computed spectrum, including many high order terms. Since in practice $I(\delta)$ is sampled, all their sampling images will be present also.

Simulation performed with a general purpose computer (Figure 33-7) shows that the first subharmonics are indeed more intense; most of the others can be treated as noise. This is even more true in the case of a complex spectrum. Thus, in general, Table 33-1 can be used to predict the additional spectrum noise. The real-time computer results from an artificial interfero-

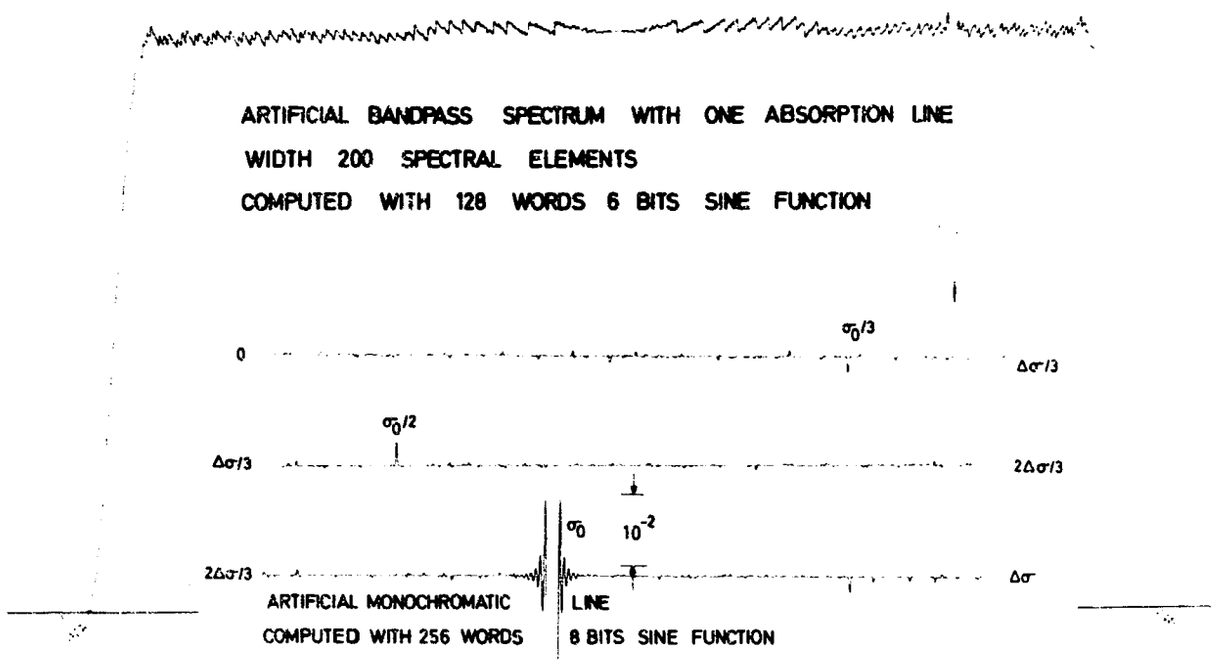


Figure 33-7. General Purpose Computer Simulation of Absorption and Emission Spectra Computed by Classical FT Using Quantized Sines

gram confirm the noise-like aspect of the errors. This can be seen by looking ahead to Figure 33-18.

33-3.3 Sine Samples Address Computation

33-3.3.1 PRINCIPLE

Let δ_1 be the interferogram sampling interval, and let $I_0, I_1, \ldots, I_n, \ldots, I_N$ be the samples; the path difference for I_n is $n\delta_1$.

Let σ_1 be the spectrum sampling interval, and $B_0, B_1, \ldots, B_m$ the samples; the wavenumber corresponding to B_m is $m\sigma_1$. In general the entire spectrum will not be computed, but only $M \leq N$ samples; if $\mu\sigma_1$ is called the lowest wavenumber in the range, the index (m) will vary from μ to $\mu+M$.

Computing a sampled spectrum* means the integral

$$B(\sigma) = \int_0^{\delta_{max}} I(\delta) \sin 2\pi\sigma\delta \, d\delta \qquad (33\text{-}12)$$

is replaced by the sum

$$B_m = \sum_{n=0}^{n=N} I_n \sin 2\pi mn\sigma_1 \, \delta_1. \qquad (33\text{-}13)$$

Since the function $z \simeq \sin(\pi/2)x$ has been stored, but only with $0<x<1$, it is necessary to find the proper addresses x_p from the argument of the sine $2\pi mn\sigma_1\delta_1$. A very simple solution is found if $\sigma_1\delta_1$ is a negative power of 2.

While δ_1 cannot be chosen at will (since it has to be a multiple of a reference wavelength λ_{ref}) † there is some freedom in selecting σ_1. One must only satisfy

$$\sigma_1 \leq 1/2\ \delta_{max} = 1/2N\ \delta_1. \qquad (33\text{-}14)$$

Consider the first integer h greater than L_2N, or, in other words, such as $2^{h-1} < N < 2^h$. Select a spectral sampling interval

$$\sigma_1 = \frac{1}{2^h\ \delta_1}, \qquad (33\text{-}15)$$

which satisfies the above condition and in general gives some oversampling. Actually there will be no oversampling at all if it is elected to take $N = 2^h$ (which

* Functions of time could of course be treated simply by replacing δ by time t and σ by frequency f.

† In the case of the 10^6 sample interferograms, δ_1 is a multiple of λ_{ref} 100 = 3.50 Å.

is no great loss of freedom), but this is not compulsory. Then

$$\sigma_1 \, \delta_1 = \frac{1}{2^h} \tag{33-16}$$

and

$$B_m = \sum_{n=0}^{n=N} I_n \sin 2\pi \frac{mn}{2^h}. \tag{33-17}$$

The number $mn/2^h$ can be either $\gtrless 1$, but in any case only the fractional part F (with $0<F<1$) is of interest. Consider F written in binary form with h bits.

$$F = \tfrac{1}{2}[(A_0 2^0 + A_1 2^{-1}) + (A_2 2^{-2} + \cdots + A_{12} 2^{-12}) + \cdots + A_h 2^{-(h-1)}], \tag{33-18}$$

in which the coefficients A are either 0 or 1. The first two terms (A_0, A_1) and the next 10 terms ($A_2 \ldots A_{12}$) have been isolated.

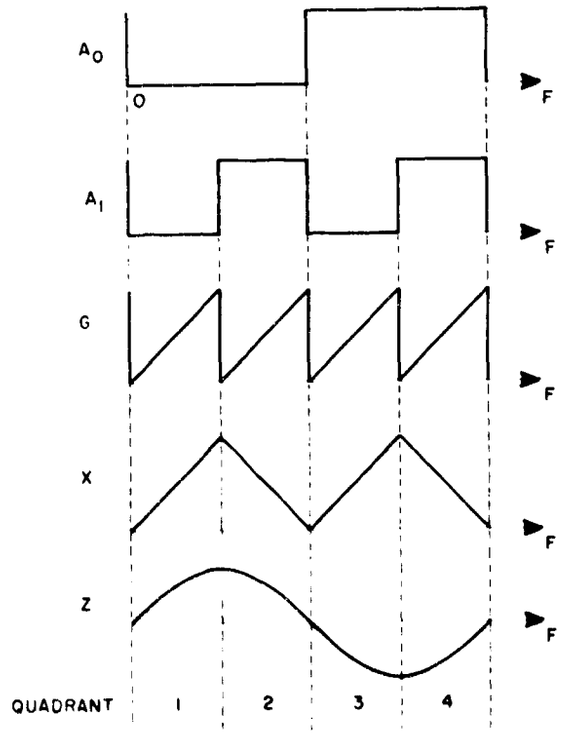


Figure 33-8. Reconstruction of a Full Sine Function Period From the Stored Quarter of a Period

The first two terms A_0, A_1 indicate in which of the four quadrants the desired sine sample lies (Figure 33-8). If $A_0=0$ (first two quadrants), the sign of the sine will be positive; if $A_0=1$ (second two quadrants), the sign is negative.

The next ten terms $A_2 \ldots A_{12}$ give the wanted address in the $1024=2^{10}$ word memory. If

$$G = 2[A_2 2^{-2} + \cdots + A_{12} 2^{-12}], \tag{33-19}$$

which gives $0<G<1$, then

$$x_p = G \quad \text{if } A_1 = 0 \quad \text{(first and third quadrants)} \tag{33-20}$$

$$x_p = 1 - G \quad \text{if } A_1 = 1 \quad \text{(second and fourth quadrants)} \tag{33-21}$$

33-3.3.2 REAL-TIME OPERATION

Computing in real time Sum (33-17) means that as soon as sample I_n becomes available the M sine samples corresponding to the given n and to $m=\mu$, $\mu+1, \mu+2, \ldots, \mu+N$ must be found. Considering the product $P=mn$ during this sequence gives

$$P = \mu n, (\mu+1)n, (\mu+2)n \ldots, (\mu+M)n \tag{33-22}$$

or, putting $j=m-\mu$ with $0<j<M$,

$$P = n\mu, n\mu+n, \ldots, n\mu+nj, \ldots, n\mu+nM. \tag{33-23}$$

Thus, it is possible to go from one P to the next by *adding* n. After the next sample I_{n+1} the sequence is the same with n increased by one unit. Figure 33-9 gives the evolution of P; it is seen that P can be obtained performing *additions* only. Figure 33-10 gives the corresponding block diagram.

The first step is the selection of the sampling interval σ_1 and of the starting point in the spectrum, which is defined by μ. In order not to have the computer handle needlessly large numbers, aliasing of the spectrum is utilized. The free spectral range $\Sigma = 1/2\delta_1$ has been selected in order to avoid overlapping. Thus, $g\Sigma<\sigma<(g+1)\Sigma$, where g is an integer. There are two cases:

33-3.3.2.1 Case Where g is Even (Figure 33-11(a))

In the range $0<\sigma<\Sigma$ one finds an identical image of the "true" spectrum. Thus, instead of using μ

n \ m	μ	μ+1	μ+2	μ+j	μ+M
1	μ	μ+1	μ+2	μ+j	μ+M
2	2μ	2μ+2	2μ+4	2μ+2j	2μ+2M
3	3μ	3μ+3	3μ+6	3μ+3j	3μ+3M
---	-----	-----	-----	-----	-----
n	nμ	nμ+n	nμ+2n	nμ+nj	nμ+nM

EVOLUTION OF P=mn → MEANS +n IN ACC. A
↓ MEANS +μ IN ACC. B

Figure 33-9. Evolution of the Product $P=mn=n\mu+\mu j$ Available in the 20Bit Parallel Adder. Each row corresponds to one input, each column to one output. Sequence of operation is the same as when reading a printed text

defined by $\mu\sigma_1=\sigma_{min}$ it is possible to use μ' such as $\mu'\sigma_1=\mu\sigma_1-g/\Sigma$, which gives:

$$\mu'=\mu-g2^h,$$

with

$$0<\mu'<N, \qquad (33\text{-}24)$$

since the free spectral range contains N spectral elements. This means μ' can be expressed in binary form with h bits, that is, stored in an h element register. When designing the computer one must consider H, the maximum possible value of h; $H=20$ has been taken.

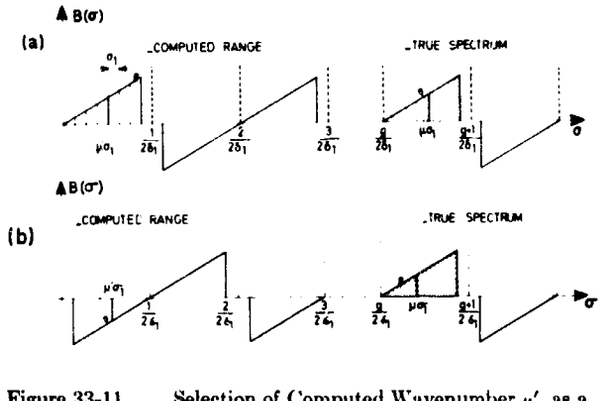


Figure 33-11. Selection of Computed Wavenumber μ', as a Function of Wanted Wavenumber μ. Figure corresponds to the sine transform of an odd interferogram

33-3.3.2.2 Case Where g is Odd (Figure 33-11(b))

In the same range $0<\sigma<\Sigma$ an inverted image of the spectrum is found. Then consider the index μ corresponding to the maximum wanted wavenumber ($\mu\sigma_1=\sigma_{max}$) and actually use μ' such that

$$\mu'\sigma_1=(g+1)/\Sigma-\mu\sigma_1, \qquad (33\text{-}25)$$

which means

$$\mu'=(g+1)2^h-\mu, \qquad (33\text{-}26)$$

and one still has $0<\mu'<N$.

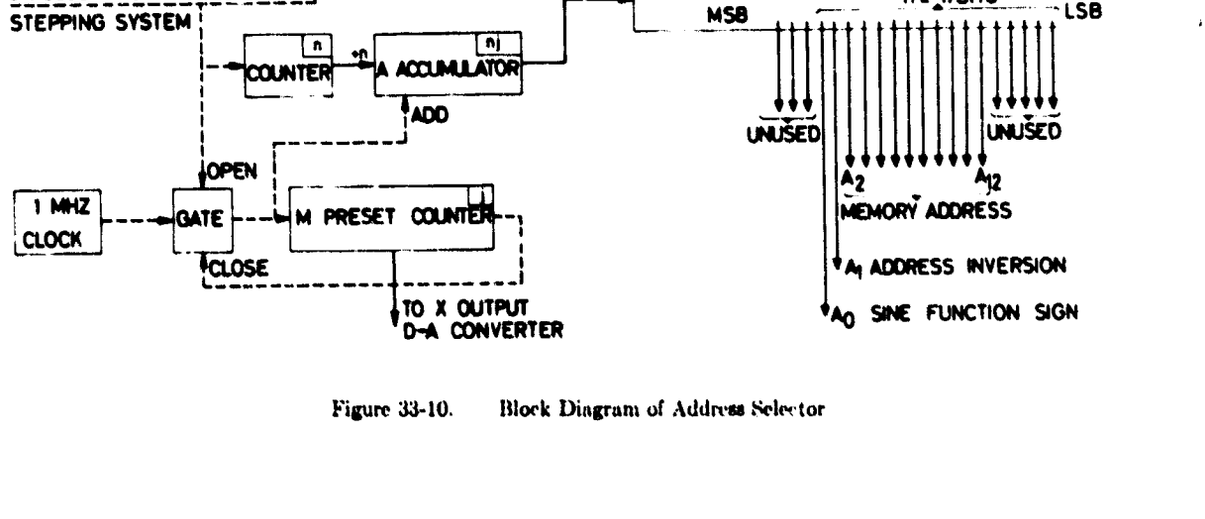


Figure 33-10. Block Diagram of Address Selector

Before the beginning of the interferogram the operator enters μ' into a manual 20 switch register. Each interferometer synchronization pulse I delivered after the recording of each sample causes the addition of the fixed register content μ' into an accumulator B; after n input samples B will hold $n\mu'$. The I pulses are directly stored in a 20 bit binary counter which registers n. Finally a gate is opened which delivers a train of M clock pulses at a 1-MHz rate; these are counted by an M preset counter which closes the gate. This counter holds number j; the contents are read by the digital-to-analog converter which gives the X oscilloscope output. Each clock pulse in the train causes the addition of n into a second accumulator A; after receiving pulse number j, A holds nj.

A parallel adder with $H=20$ bits sums the outputs of both accumulators, and thus gives the sum $n\mu'+nj$ automatically clipped to H bits. Since only $h \leq H$ bits are needed to represent in binary form F, the fractional part of $P=nm$, the operator selects the h least significant bits with a multipole switch. In this way the wanted spectral sampling interval $\sigma_1 = 1/2^h \delta_1$ is entered into the computer.

Out of these h bits only the 12 most significant are actually used. The first is A_0 and gives the sine sample sign; the second, A_1, inverts the order of the sine memory addresses; the next ten give the addresses. The remaining unused LSB would correspond to a finer subdivision of the sine function than is available (They would be needed if the two-stage addressing scheme mentioned in Section 33-3.2.2 were used), and these LSB are not used for addressing. However, they are necessary for correctly storing P.

33-3.4 Data Handling

The multiplier has only a 10 bit capacity, which is quite insufficient to handle the interferogram. The technique already described[1] for recording large dynamic range interferograms involves manually increasing the gain by factors of 10, 10^2, and 10^3 after the signal has dropped. The gain changes are recorded by a special character on the digital magnetic tape for correction during the *a posteriori* computation.

The same procedure has been kept for analysis by the real-time computer but the selected gains are now 1, 4, 16, 64, 256, $1024 = 2^{10}$. The 10 bit multiplier output is correspondingly shifted by 0, 2, 4, 6, 8 or 10 bits at the 24 bit adder input. This is achieved by a direct link between the recording system manual attenuator and the adder.

The integrating digital voltmeter consists of a voltage-to-frequency converter (100 kHz maximum, giving 2000 counts maximum at 50 samples per sec and more at slower recording rates) and two reversible counters. The first counter is decimal and provides 5 decades BCD output to the tape recorder. The

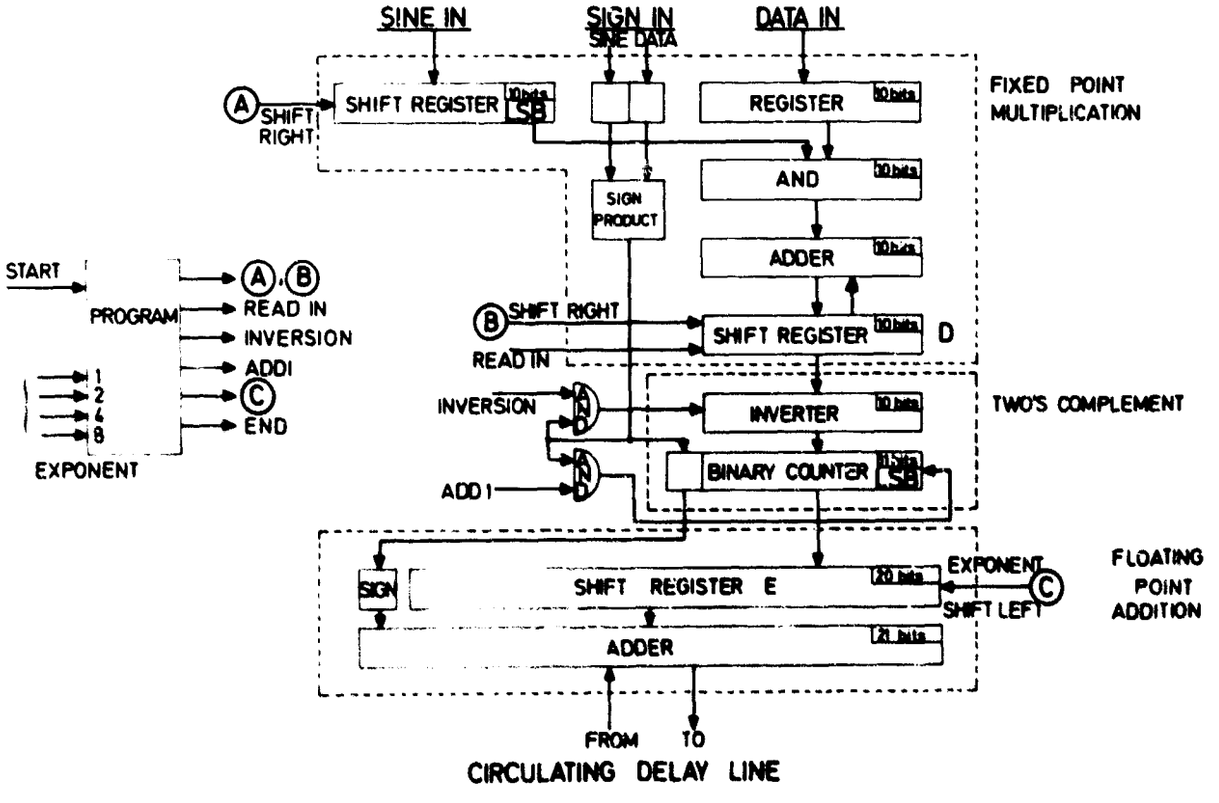


Figure 33-12. Block diagram of Arithmetic Unit. The program generator contains a synchronizer and a preset synchronous counter operating from a 15.4 MHz oscillator. It produces the SHIFT and READ IN pulses. Another preset device also produces SHIFT pulses, according to the exponent value. These pulses are 10 nsec wide with a repetition period of 15 nsec

second is binary (14 bits + sign) and the output goes to the real-time computer. The counter automatically detects and sends the 10 most significant bits used to the multiplier input. The sign and exponent, which can be 2^0, 2^1, 2^2, 2^3, or 2^4 are sent to the adder. The net result is that any interferogram sample between ± 32.676 can be measured with 10 bit accuracy, making less critical the gain adjustments when starting an interferogram. The exponent is added to the manually selected exponent, which gives an overall dynamic range of 24 bits.

33-4 SUMMARY CIRCUIT DESCRIPTION

33-4.1 Arithmetic Unit (Figures 33-12 and 33-13)

The arithmetic unit contains a multiplier, a two's complement circuit, an adder, and a program generator which delivers pulses to command the sequence of an elementary computation cycle. The operations are performed using a pure binary two's complement code. The main features are: (a) fixed point multiplication of the two 10 bits + sign words; (b) two's complement of the multiplication result; and (c) floating point addition of 10 bit + sign words.

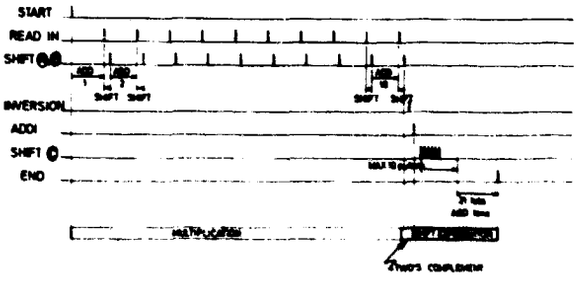


Figure 33-13. Timing of Arithmetic Unit

The principle of operation is briefly described as follows:

33-4.1.1 MULTIPLICATION

The technique is classical, and uses parallel addition and shifting. The DATA multiplicand and the SINE multiplier are respectively read into a register and a shift register; initially the shift register D is cleared.

In the first step the multiplicand is multiplied by the multiplier LSB; in binary this is a simple AND operation. The result is added to the register D content. After a delay which permits carries to propagate, the adder output is read into register *D* upon a READ IN pulse command. Then a pulse fed into *A* and *B* shifts one place to the right the contents of the registers.

After ten similar steps register D contains the result of the multiplication limited to the 10 MSB.

33-4.1.2 TWO'S COMPLEMENT

If the sign of the product is minus it is necessary to take the two's complement of register D content. This is done in two steps:

(1) inversion and transfer into a binary counter, and

(2) a pulse ADD 1 is fed into the counter.

If the sign is plus, the content of D is transferred directly into the adder.

33-4.1.3 ADDITION

Exponent pulses are fed into the C input of the shift left register, the output of which is parallel added to the word coming out of the delay line memory. The result is recirculated into the memory.

The arithmetic unit is implemented with fast ECL logic circuits (Motorola MECL II). The total time for a computation cycle is actually 950 nsec; multiplication takes 650 nsec, addition 520 nsec, two's complementing and various transfers 50 nsec.

Using the same circuits but a more elaborate design including a parallel multiplier, a look ahead carry type adder, and a data selector multiplexer in place of the shift register, the time could be reduced to less than 100 nsec.

33-4.2 Output Memory

This computer was designed to work with any sequential access memory. Magnetostrictive delay lines appeared best for a first try; MOS shift registers were also considered.

The lines (Figure 33-14) are monitored by a pulse generator, which includes a quartz oscillator of period T delivering clock pulses CLK. A modulo M divider gives out FLAG pulses and M-3 pulses which correspond to the M-3 state of the counter. A synchronization circuit is also needed. If D is called the line delay time, obviously D = MT.

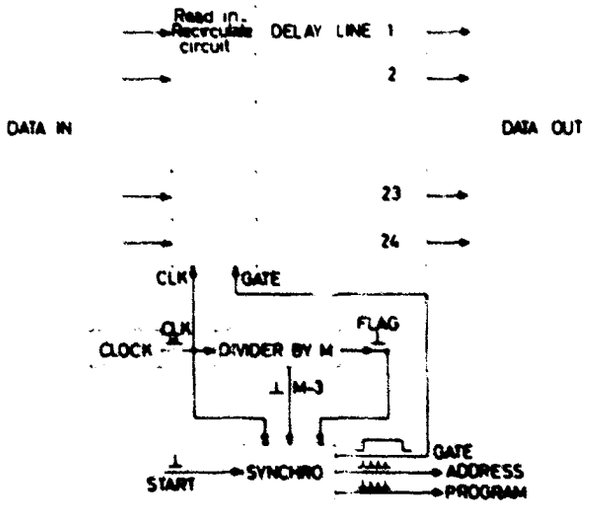


Figure 33-14. Output Memory Block Diagram. The GATE trailing edge is used to inhibit one CLK pulse at M counter input; the corresponding circuitry is not represented

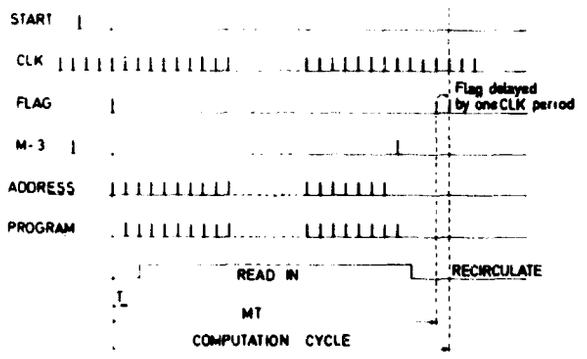


Figure 33-15. Timing of Output Memory

A START pulse (Figure 33-15) given by the DVM triggers the synchronization circuit, which produces pulses to command M-3 elementary cycles; this means M-3 ADDRESS pulses, M-3 PROGRAM pulses, and a GATE. The GATE enables writing the results into the delay lines and permits their recirculation at the end of the cycle.

All these operations are illustrated in Figure 33-16. After each input sample is treated, note that the memory content is shifted by one T period relative to FLAG. Thus, before triggering a new cycle it is necessary to delay the FLAG; this is done at the end of a cycle by inhibiting one CLK pulse at the M divider input.

During each elementary cycle the following operations proceed simultaneously: (a) Word j read out, (b) Word j+1 address generation and sine extraction, (c) Word j program execution; and (d) Word j−1 read on. The actual elementary cycle time is less than 1 μsec.

By taking a delay line with D=20 msec, which is the maximum delay commercially available, and T=1 μsec, it is possible to store 20 000 words; if 24 parallel lines are used each word has 24 bits.

A direct oscilloscope output with a 50-Hz sweep repetition rate is obtained through an 11 bit + sign D→A converter; the 11 MS actually used bits can be selected. A pen recorder output is produced by selecting one sample per line recirculation. Thus, the spectrum is scanned at a 50 sample per sec rate.

33-5 OUTPUT INTERPOLATOR-APODIZER

The digital computer output consists of spectral samples with separation σ_1. To obtain a sufficiently accurate spectral curve two techniques are possible. One could use considerable oversampling ($\sigma_1 \leq (1/5)\cdot(1/2\delta_{max})$) which would mean a proportional reduction of the spectral range for a given spectrum storage capacity. A better solution is the use of an output interpolation. J. Connes (Chapter 6) shows that convolving the output samples by the instrumental line shape (ILS) provides the wanted continuous spectral curve, that is, both apodization and interpolation at the same time. Between 2 "primary" adjacent spectral samples, 5 "secondary" interpolated samples are computed; 20 primary samples and 10 samples of the (symmetrical) convolving function are needed, 20 multiplications and 20 additions being involved. Extremely good accuracy in the profile of the final ILS is realized.

33-5.1 Digital Interpolator

A fully digital interpolator-apodizer could be built. The wanted ILS would be stored (5 sets of 10 samples). Since 20 multiplications and additions are needed for each secondary sample and 5 secondary samples are wanted per primary sample, 100 multiplications and additions take place for each primary sample. In order not to slow down the FT computer, only 1/100 of the primary spectrum storage should be read during each 20 msec recirculation. Presentation of the complete spectrum would take 2 sec, which is more than adequate.

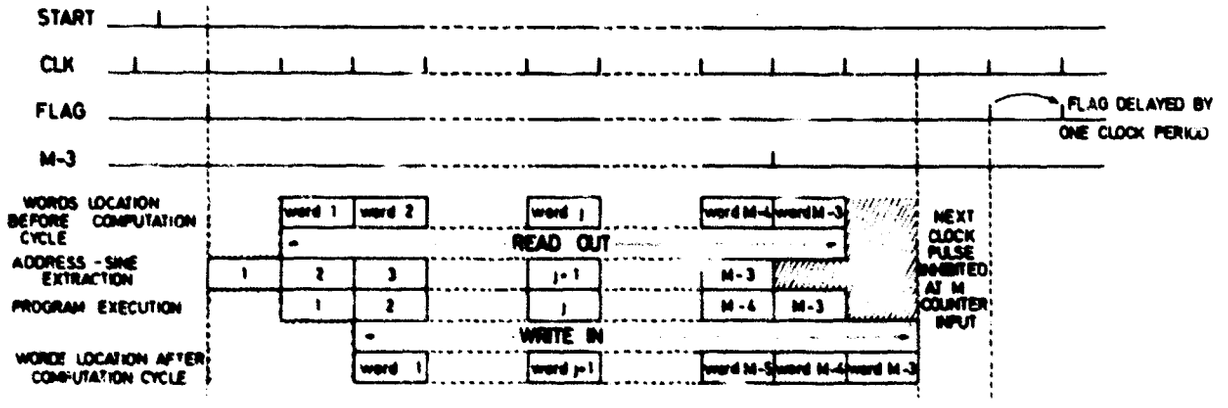


Figure 33-16. Computation Cycle Sequence. Words location into the memory before and after computation cycle is shown. The output memory has M words capacity, of which M−3 are actually used for storing the spectrum; the remaining 3 give the zero base line

If the separation between the selected primary samples is taken equal to σ_1, apodization is optimum when the interferogram scan has reached δ_{max}. However, by selecting $2\sigma_1, 3\sigma_1, \ldots$ intervals, one can get the same apodization at $\delta_{max}/2, \delta_{max}/3, \ldots$.

The system would include a very small read-only memory for storing one (or several) convolving function, a random access memory with M/100 word capacity for storing after each recirculation the needed primary samples, a multiplier and adder. Since it is not necessary to have continuous apodized and interpolated presentation of the data, one could use the already available multiplier and adder by interrupting the interferogram recording for 2 sec at suitable intervals.

33–5.2 Analog Interpolator

It is hoped that a much simpler all-analog interpolator will prove accurate enough. The system is not fully tested and only the principles will be given. See Figure 33-17.

Since the delay-line memory is clocked by a quartz crystal oscillator with frequency $\Phi = 1$ MHz, the primary spectral samples, after D→A conversion, are available as accurately timed voltage pulses. Their width is constant and small compared with their 1 μsec separation. If the wanted ILS is $S(\sigma)$, the necessary convolution can be performed by sending the pulses through a linear filter with impulse response $S(\Phi\sigma, t)$. This filter can be understood as performing simultaneously the three operations necessary to a digital convolution: storage of the primary samples, weighting by suitable coefficients and addition of the results. If 20 samples must be used, the total duration of the impulse response is 20 μsec.

A filter with any given impulse response can be approximated by a multitapped analog delay line, the outputs of which are summed after suitable weighting. Since higher relative accuracy is needed for the central peak than for the much weaker sidelobes, two separate delay lines B and C will be used, each with 10 taps. The first, with 20 μsec total delay, will synthesize the sidelobes and the second, with 2 μsec delay, will synthesize the central peak.

The ILS can be modified by changing the array of weighting resistors. A change of delay lines could in principle give optimum apodization for several path differences, but this seems hardly necessary.

Better approximation for a given number of taps is realized if the input pulses are not rectangular but Gaussian. These can be approximated in various manners, for instance by applying the same principle again in a third delay line A.

33–6 RESULTS

The instrument has not yet reached the final anticipated form described here. The multiplier, adder, and address selector are complete and the system operates at the normal clock rate of 1 MHz. However, an available 8 bit, 256 word read and write memory has been used for storing the sines, while waiting for delivery of the 1000 word, 10 bit read only final version. Only one delay line has been installed. The words are stored in serial form and contain only 20 bits (instead of 24 bits). Capacity is thus limited to 1000 spectral samples. Completion of the system means an increase in capacity only, not in speed, and involves no basic or technical problems.

The accuracy of the computer has been tested with an artificial sampled sine wave interferogram. Figure 33-18 shows four portions of the computer ILS (unapodized) and shows the errors implied by the use of a quantified sine function. The input is an artificial sampled sine wave $(0, +1, 0, -1, 0 \ldots)$. Each of the four traces corresponds to the full 1000 sample output and individual samples are not visible. Neither apodization nor interpolation are used. The central peak of the sinc instrumental function is shown, together with background regions. Major oscillations are distant secondary maxima of the sinc function. "Noise" is due to sine quantization (8 bits, 256 words). When the number of input samples increases, the central peak and the noise grow at the same rate, that is, relative noise intensity is constant. Unlike true noise originating in the interferogram, this noise shows no correlation between samples, or in other words is not smoothed by the instrumental line shape.

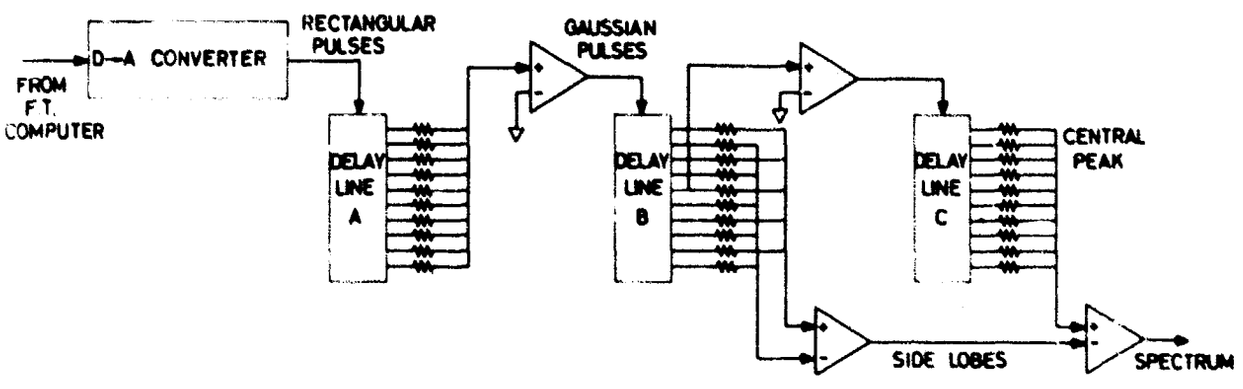


Figure 33-17. Analog Interpolator-Apodizer. One of the outputs of delay line B is taken as input of C

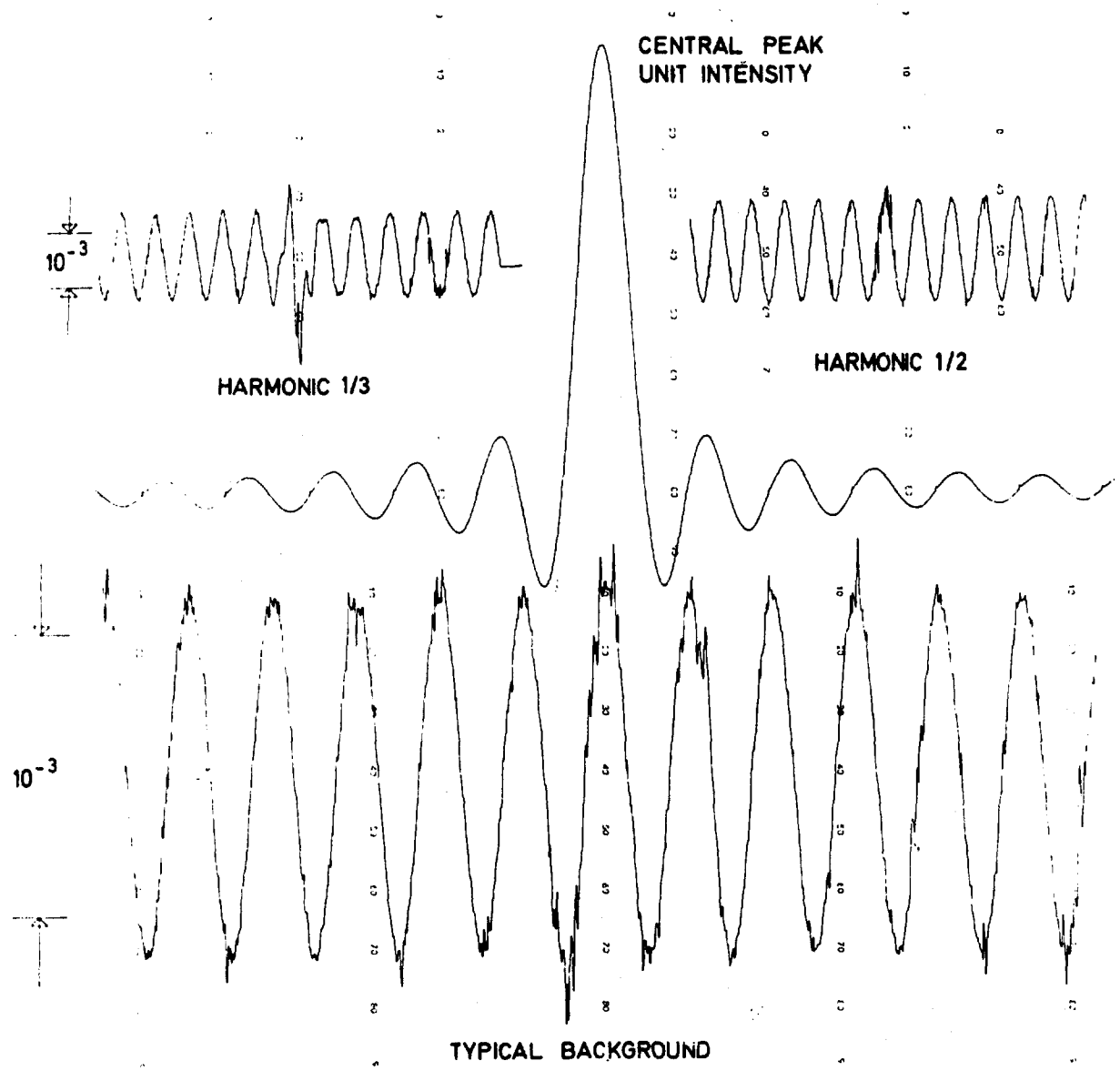


Figure 33-18. Accuracy Test of Real-Time Computer and Digital-to-Analog Output Converter

Figure 33-19 shows the expected sharpening of the sinc function ILS when the interferogram length increases.

Figure 33-20 illustrates how the computer can be used for either high or low resolution problems. The absorption spectrum of N_2O between 4000 and 5000 cm^{-1} is presented, and is explained as follows. A spectral region from 3800 to 5300 cm^{-1} is selected with an interference filter. Each set of curves corresponds to one interferogram and one computer setting. Within each set resolution increases by factors of two. The full 1000 sample output is presented. No apodization nor interpolation are used. The pen recorder output is shown.

(a) Interferogram sampling interval $\delta_1 = 0.175\,\mu$, computed range $M\sigma_1 = 3300\ cm^{-1}$. The number of input samples N is indicated. The maximum path difference $\delta_{max} = 0.22$ cm.

$\delta_1 = 0.175\,\mu$ $M\sigma_1 = 3300\ cm^{-1}$ $\delta_{max} = 0.22$ cm

(b) $\delta_1 = 2.7\,\mu$ $M\sigma_1 = 900\ cm^{-1}$ $\delta_{max} = 1.7$ cm

(c) $\delta_1 = 2.7\,\mu$ $M\sigma_1 = 55\ cm^{-1}$ $\delta_{max} = 4.3$ cm

(d) $\delta_1 = 2.7\,\mu$ $M\sigma_1 = 7\ cm^{-1}$ $\delta_{max} = 86$ cm

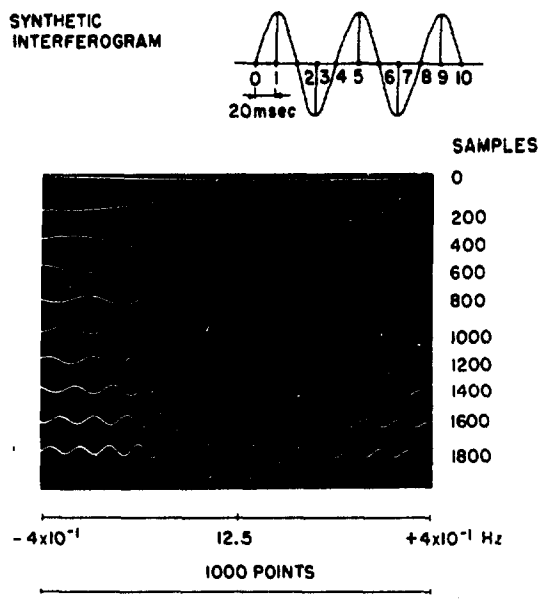


Figure 33-19. Sharpening of ILS as the Number of Input Samples is Increased (Oscilloscope Output). The synthetic interferogram is the same as for Figure 33-17

Figure 33-21 gives the first results of the incomplete interpolator. Only delay line C is used, giving synthesis of the central peak alone. The input "Gaussian" pulses are crudely approximated by filtering rectangular pulses with an RC network. A symmetrical, but yet unapodized, instrumental line shape is thus realized. The improvement compared to a simple RC smoothing is already apparent.

33-7 CONCLUSIONS

The computer described was built for Fourier spectroscopy as practiced at Laboratoire Aimé Cotton, that is, for very high resolution in the near infrared with checking of only a small fraction of the entire spectrum. However, the computer might be used for any problem in which the speed and capacity limitations given in Figures 33-1 and 33-2 would not be exceeded.

The present performance could definitely be improved. Using commercially available integrated circuits and memories only, a factor of at least 3 (possibly 10) improvement in speed could be realized, giving a basic computational time of

$$1\times10^{-7}\ \text{sec}<\theta<3\times10^{-7}\ \text{sec}.$$

This increase in speed would make the instrument competitive with the time compression analyzers

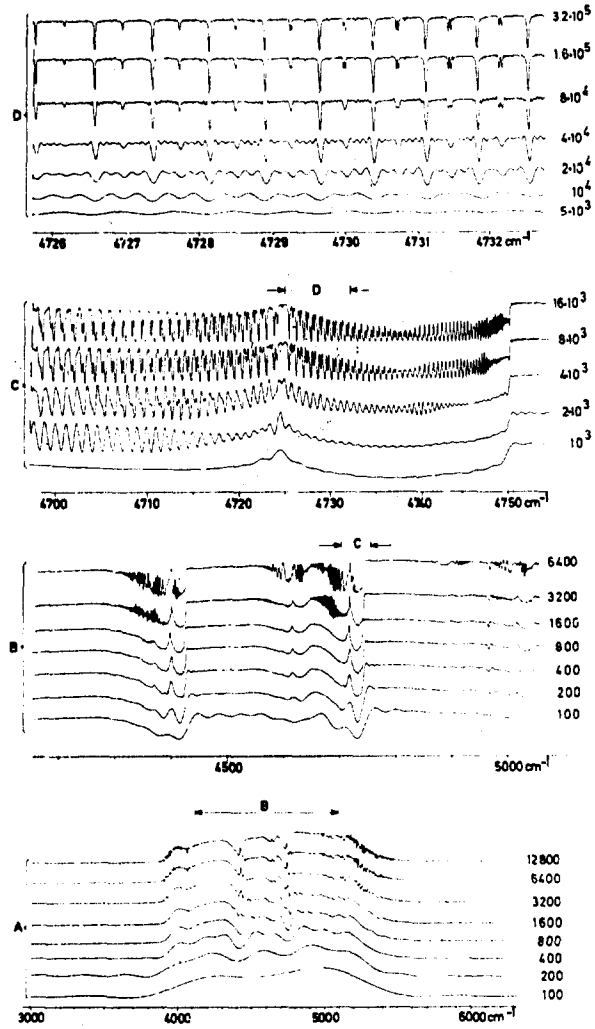


Figure 33-20. Absorption Spectrum of N_2O

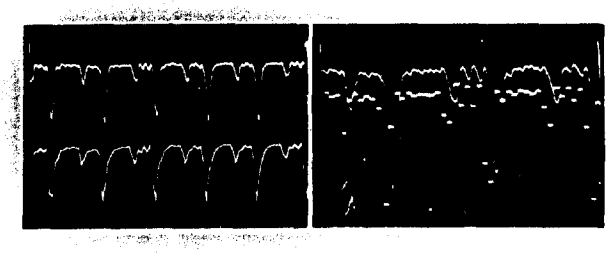


Figure 33-21. N_2O Lines (Same Region as Figure 33-20d) With and Without Analog Interpolator (Oscilloscope Output). The device is not completed and interpolates but does not apodize (see text). Figure 33-20(a) compares the output with the sampled input and Figure 33-20(b) compares the output with an RC smoothing filter

which have many applications in the fields of acoustical frequencies. The advantages of better accuracy through all digital operation, and of unlimited resolution because of the spectrum storage principle, would be kept.

The 20 000 words (and 500 000 bits) output capacity selected as a goal roughly corresponds to the economic limit of magnetostrictive delay line storage. However, the instrument could, with very little modification, be adapted to drum or disk operation. These provide the cheapest storage per bit above 10^6 bits.

Thus it is apparent that real-time computation of the entire spectrum in almost all Fourier spectroscopy problems is technically feasible with a system of this type. The only exception is Fourier spectroscopy of high intensity but very short life time sources. In such cases the input sampling rate should be limited by receiver bandwidth only, and might exceed the computer speed, at least when a large M is wanted. See Figure 33-2. In such a case, *a posteriori* computation by FFT remains the only choice.

On the other hand, simplified versions of the computer described here, with reduced capacity and/or speed and accuracy could also be built. In all cases the system should be much cheaper than the equivalent FFT computer due to the basic simplicity of the addressing scheme and to the elimination of large random access memories.

Furthermore, the majority of Fourier spectroscopists use continuous scans and considerable oversampling (N≫M); in this case storing the spectrum instead of the interferogram gives a reduction of storage capacity. Thus even when recording time is short, and real-time computation is of no interest by itself, the computer described here provides the simplest solution.

Acknowledgments

The authors wish to thank H. Delouis for programming the computer simulation of the quantified sine transforms. The success of the computer is due to the constant assistance of J. Seguin, who was largely responsible for the addressing system and the pen recorder output, and to the careful wiring and assembling work of H. Durand and J. Dessens.

References

1. Connes, J., Delouis, H., Connes, P., Guelachvili, G., Maillard, J. P., and Michel, G. (1970) *Nouv. Rev. Opt. App.* **1**: 3.

2. Hoffman, J. E. (1968) AFCRL Instrumentation Paper No. 146.

3. Hoffman, J. E., and Vanasse, G. (1966) *Appl. Opt.* 1167.

4. Hoffman, J. E., and Vanasse, G. (1967) *J. Phys.* **28**, C2: 79.

5. Gayles, J. N., Honzik, W. L., and Wilson, D. O. (1970) *IBM J. Res. Develop.* **14**: 25.

6. Cole, T. W. (1968) *Optics Techn.* **31**.

7. Bickel, H. J. (1961) U.S. Patent 3.013.209.

8. Fellgett, P. B. (1967) *J. Phys.* **28**, C2: 53.

9. Michelson, A. A., and Stratton, S. W. (1898) *Am. J. Sci.* **5**: 1.

10. Strong, J. D., and Vanasse, G. (1958) *J. Phys.* **19**: 192.

11. Yoshinaga, H., Fujita, S., Minami, S., Suemoto, Y., Inoue, M., Chiba, K., Nakano, K., Yoshida, S., and Sugimori, H. (1966) *Appl. Opt.* **5**: 1159.

12. Ridyard, J. N. A. (1967) *J. Phys.* **28**, C2: 62.

Contents

34. Real-Time Far-Infrared Fourier Spectroscopy Using a Small Digital Computer

F. Levy and R. C. Milward
Société Coderg, 15, Impasse Barbier, 92 Clichy, France
and
S. Bras and R. Le Toullec
Laboratoire Physique des Solides, Faculté des Sciences Paris, France

Abstract

A 4096 word memory general purpose digital computer has been successfully programmed for real-time Fourier transform computations. The computer operates on-line with a double-beam far-infrared Michelson interferometer, and can transform up to 1500 frequency output points from each interferogram value for the case of a single spectrum, or up to 750 frequency points if two spectra are to be stored and ratioed. The total time of computation for 1500 frequency points is about 1.9 seconds. The performance of the interferometer computer system for spectroscopy in the 10- to 1000-cm^{-1} spectral region is illustrated and discussed.

34–1 INTRODUCTION

The principal disadvantage of the interferometric Fourier transform technique for infrared spectroscopy has been that the spectrum of interest is not immediately observable but has to be recovered from a lengthy Fourier transform computation of a recorded interferogram function. This means that in practice there is often no facility to judge the quality of the spectrum during an interferogram scan, and the rate at which spectra are eventually obtained is determined by the rapidity with which interferogram data tapes or cards can be processed by a large digital computer. This often involves a one- or two-day wait unless time sharing facilities can be arranged.

In order to obviate some of the inconveniences, various special-purpose Fourier-transform computers have been constructed to work on-line with laboratory interferometric spectrometers. Many of these devices, however, are analog[1–3] or analog-digital[4] in nature and have been restricted by the absence of digital accuracy or by the limited number of input-output values that they can process, and also by their inability to display the spectrum until the interferogram scan has been terminated. Notable exceptions have been the special-purpose digital computer of Yoshinaga et al.[5] and the analog computer of Hoffman[6] which calculate and display the spectrum after each successive interferogram data point and are capable of processing up to 1000 frequency points in the output spectrum with an intensity accuracy of better than two percent full scale.

The purpose of this paper is to describe the successful utilization of a small (4K memory) general-purpose

digital computer* for on-line Fourier spectroscopy. This computer incorporates the advantage of a real-time mode of operation, high digital accuracy, and the ability to compute up to 1500 frequency points for a single spectrum, or 750 frequency points if a ratio of two spectra is taken. The method of programming and performance of this computer when used on-line with a double-beam far-infrared Michelson interferometer is illustrated and discussed in the following sections.

34–2 INSTRUMENTATION

34–2.1 Interferometer

The Michelson interferometer was designed for routine spectroscopy in the 10- to 1000-cm^{-1} spectral region, using an aperiodic mode of scanning interferograms. The instrument is comprised of separate interferometer and sample chamber modules and employs an all-reflecting optical system of approximately f/2 aperture and 7 cm nominal beam diameter, the ray diagram of which is shown in Figure 34-1. The radiation source is a water-cooled 125-W high-pressure mercury arc lamp, and polyethylene terephthalate films of various thicknesses mounted on a rapid-change facility are used as beamsplitters. The moving Michelson mirror is mounted on a high-precision kinematic table which is driven in continuous fashion by a synchronous motor via a gear box and lead screw. The carriage displacement is measured by a moiré fringe system which generates a pulse for a minimum displacement of 2.5 μ or an optical path difference of 5 μ. The maximum drive displacement is 5 cm (a later version uses a 10-cm displacement carriage†). In order to arrange the coincidence of a moiré fringe exactly with the zero path difference

* Varian Data Machines 620/i Computer.

† Fourierspec 2000 Interferometer, Société Coderg, Paris, France.

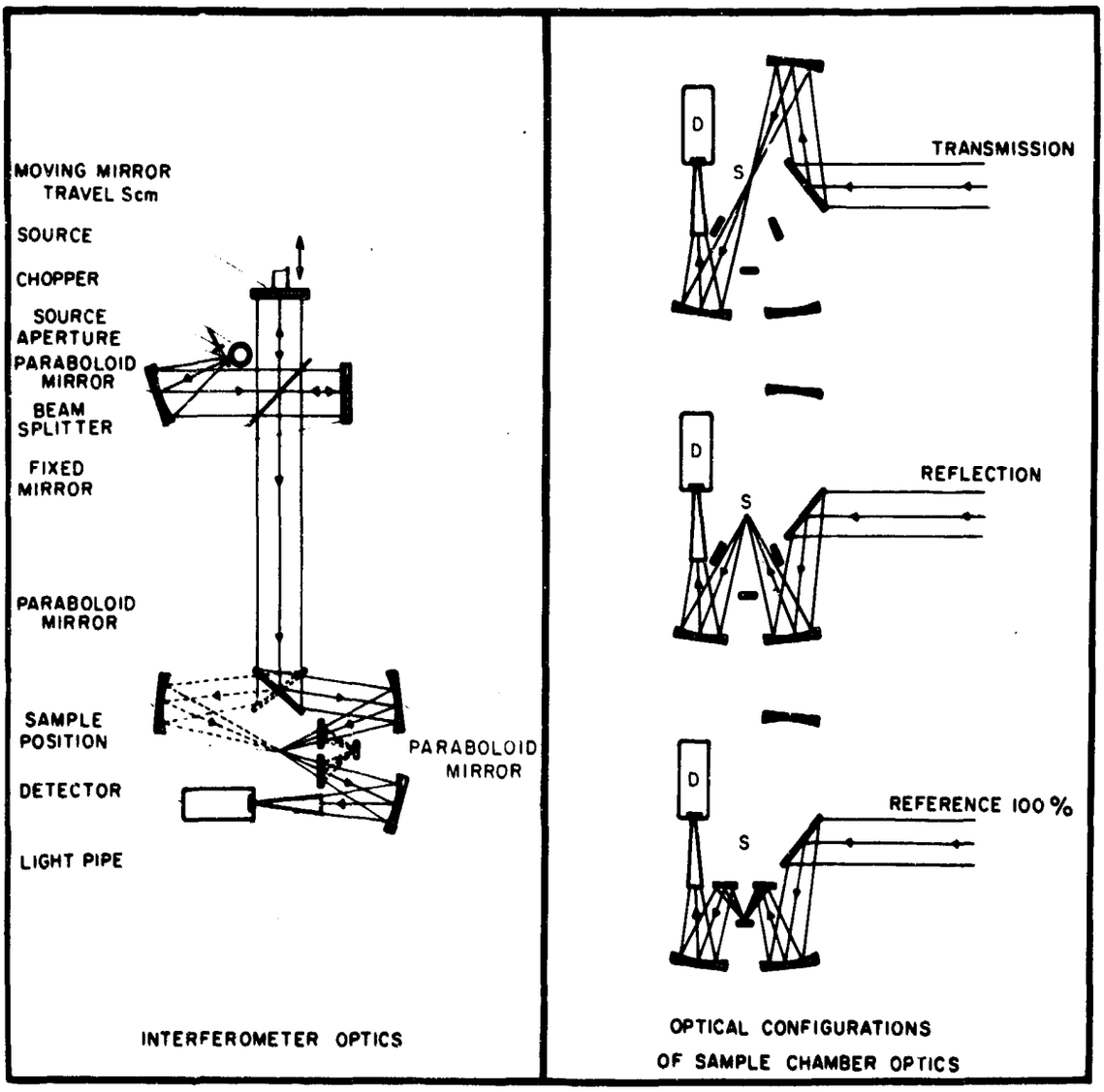


Figure 34-1. Optical Ray Diagram for the Far-Infrared Interferometer Described in the Text

interferogram position, the moiré grating can be laterally displaced with respect to the interferometer carriage, by a fine adjustment facility.

The sample chamber can be evacuated independently of the interferometer chamber by means of a high density polyethylene or "TPX"* window.[7] As shown in Figure 34-1, the sample chamber optics incorporates a system of motorized hinged mirrors which allow the sample beam to be switched through separate optical paths to enable transmission, reflection, and reference interferograms to be sequentially recorded for the same sample without disturbing the state of instrumental vacuum. The radiation is finally condensed to a 3-mm diamond-windowed Golay detector by means of internal reflection in a tapered copper light pipe. Unwanted high-frequency radiation is removed by a black polyethylene filter and other well-known far-infrared transmission filters or restrahlen reflection filters as required.

34-2.2 Electronics and Computer System

The interferometer electronic system is shown schematically in Figure 34-2. The main sequence of operations is as follows. The modulated detector output at 12.5 Hz is amplified, synchronously rectified and smoothed, and fed to a digital voltmeter which, on receipt of a moiré command pulse, converts the interferogram signal into numbers in the range ± 9999. For the purpose of real-time Fourier computations (see Section 34-3), and to increase the dynamic range of the digitizing system, the electrical interferogram signal is backed-off by a highly stable constant voltage so that the average value of the interferogram function well away from zero path is set as precisely as possible at "zero" volts. The digitized interferogram values are normally coded and transferred to punched paper tape, so that they can be processed on a large digital computer. In the system described here, however, the digitized interferogram values are fed, via a suitable interface, directly to a small Varian 620/i digital computer, where they are processed in the manner described in Section 34-3. After the spectral values have been computed and stored in the computer memory, they are displayed on a cathode ray screen. A permanent record is plotted on a chart recorder when the observed spectrum has attained the desired quality.

* Registered trademark, I.C.I. Ltd, England.

34-3 DETAILS OF COMPUTATION

The method of computation closely follows that first outlined by Yoshinaga et al.[5] and only the important details relevant to the use of the small digital computer need be described here.

The optical power spectrum $G(\nu i)$ at any wavenumber νi is computed from the single-sided oscillatory interferogram function $I'(n\,\Delta x) = I(n\,\Delta x) - \frac{1}{2}I(0)$, by the cosine Fourier transform

$$G(\nu i) = \sum_{n=0}^{n=N} I\,(n\,\Delta x)\cos(2\pi \nu i n\,\Delta x), \qquad (34\text{-}1)$$

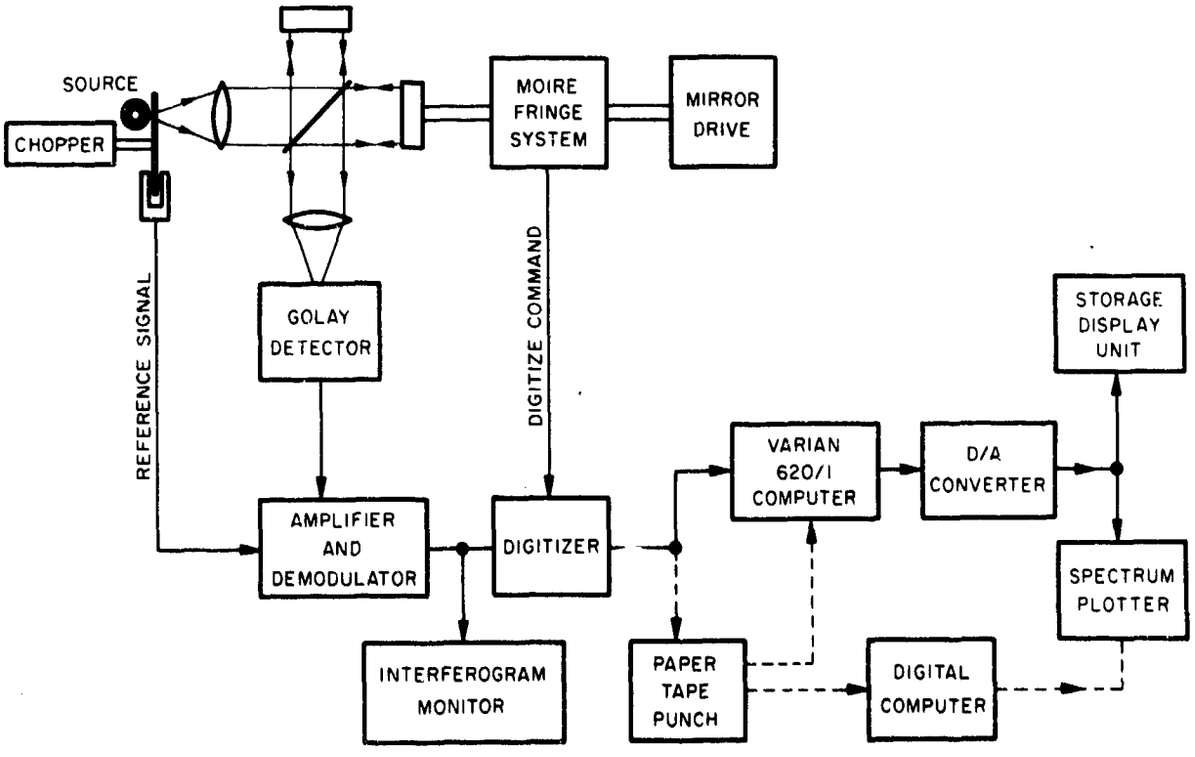


Figure 34-2. Interferometer Electronics and Computing System

where the interferogram is assumed to have been correctly sampled at regular intervals Δx, starting from the true zero path interferogram maximum and proceeding to some maximum value of optical path difference $N\,\Delta x$. The sampling interval must lie within the maximum value dictated by information theory, according to the bandwidth of optical frequencies received by the detector.

In the real-time method of Fourier transform computation, the spectrum values $G(\nu i)$ are calculated for a predetermined regular array of frequencies νi at each interferogram sample point, starting at $n=0$. The computed spectrum is then built up in detail by the addition of $G(\nu i)$ terms calculated at each successive interferogram sample point, that is, for all values of νi,

$$G(\nu i)_n = G(\nu i)_{n-1} + I'(n\Delta x)\cos(2\pi\nu i n\,\Delta x), \qquad (34\text{-}2)$$

where $G(\nu)_{-1} = 0$.

The sequence of operations used with the small digital computer system described here is illustrated in Figure 34-3 and may be summarized as follows:

(1) The computer is first initialized with the program and the appropriate experimental parameters for the Fourier transform computation. The programs used here demand that the frequency range be specified by the starting frequency ν_0, the regular interval at which the frequency array is calculated $\Delta\nu$ (in this case expressed in integer units of 0.05 cm^{-1}), and the total number of spectrum points M. The interferogram sampling interval Δx, the total number of interferogram input points N, and the choice of which one of ten different apodizing functions, or no apodization, is to be used must also be specified. Furthermore, information as to whether a single-beam or double-beam ratioed spectrum is to be calculated must be provided.

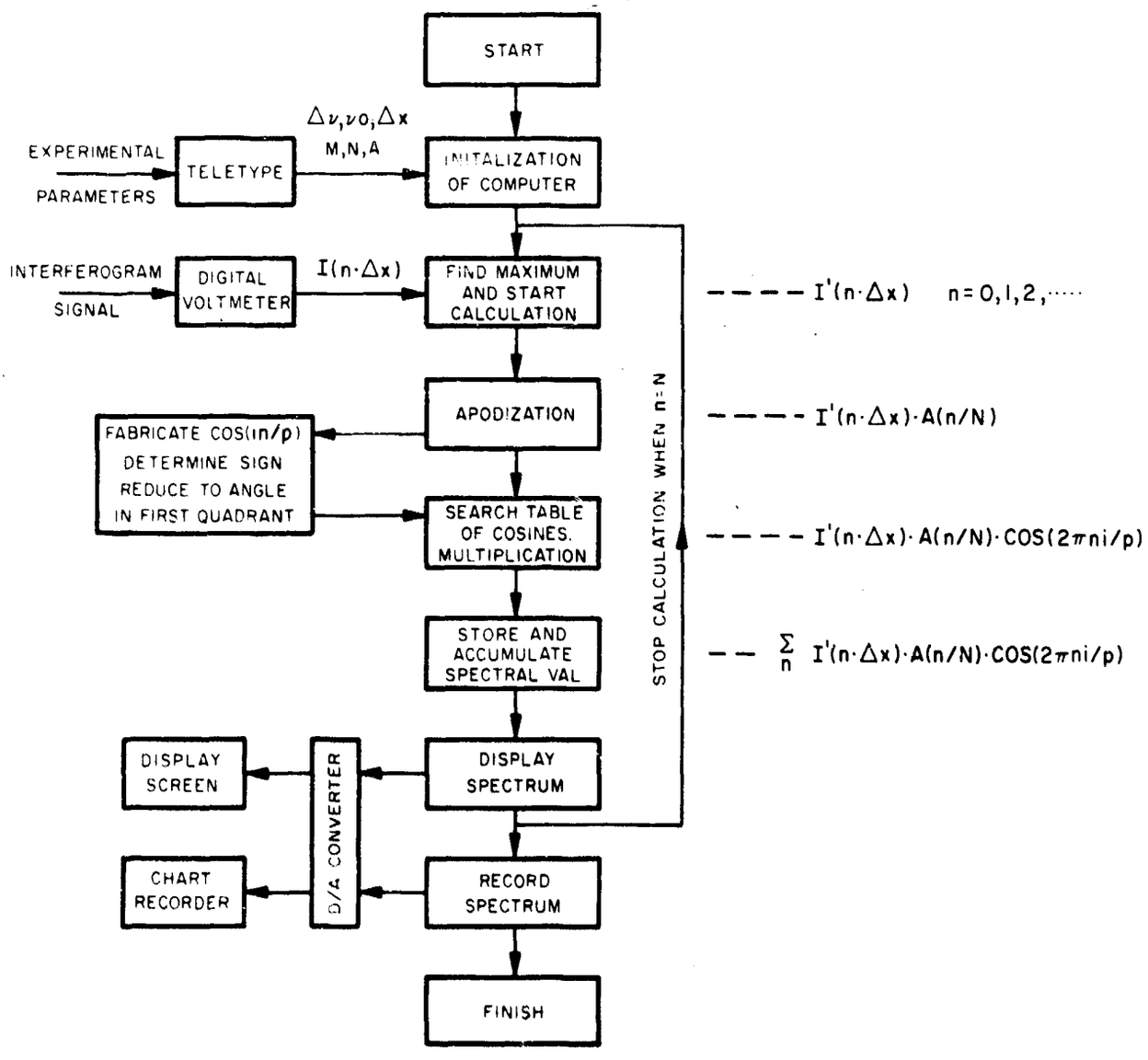


Figure 34-3. Summary of Real-Time Method of Fourier Transformation Described in the Text

(2) Values of the oscillatory interferogram function $I'(n\,\Delta x)$ are obtained directly from the voltmeter output, since the total interferogram signal has been electrically backed off so that its average value away from zero path falls at zero volts. An interferogram scan is started a short distance on the other side of zero path and the computer is instructed to commence Fourier transformations at the largest interferogram value received and to erase all preceding calculations. In this way, Fourier transformations commence at the true $I(0)$ position, provided that the moiré fringe system has been regulated to give a sample point exactly at the interferogram maximum.

(3) If required, interferogram values $I'(n\,\Delta x)$ may then be multiplied by a linear apodizing function of the form

$$A_q(n\ N) = [1 - (nq\ 10N)]. \qquad (34\text{-}3)$$

where the integer q is specified by the operator and can take values between 1 and 10. No apodization corresponds to the case $q = 0$.

(4) The values of cosine $(2\pi vin\ \Delta x)$ are determined. This is achieved in the following way. Since the frequency spectrum is calculated at regular intervals $\Delta\nu$, the cosine term may be expressed as

$$\cos(2\pi vin\ \Delta x) = \cos(2\pi i\ \Delta\nu n\ \Delta x) = \cos(2\pi in\ p), \qquad (34\text{-}4)$$

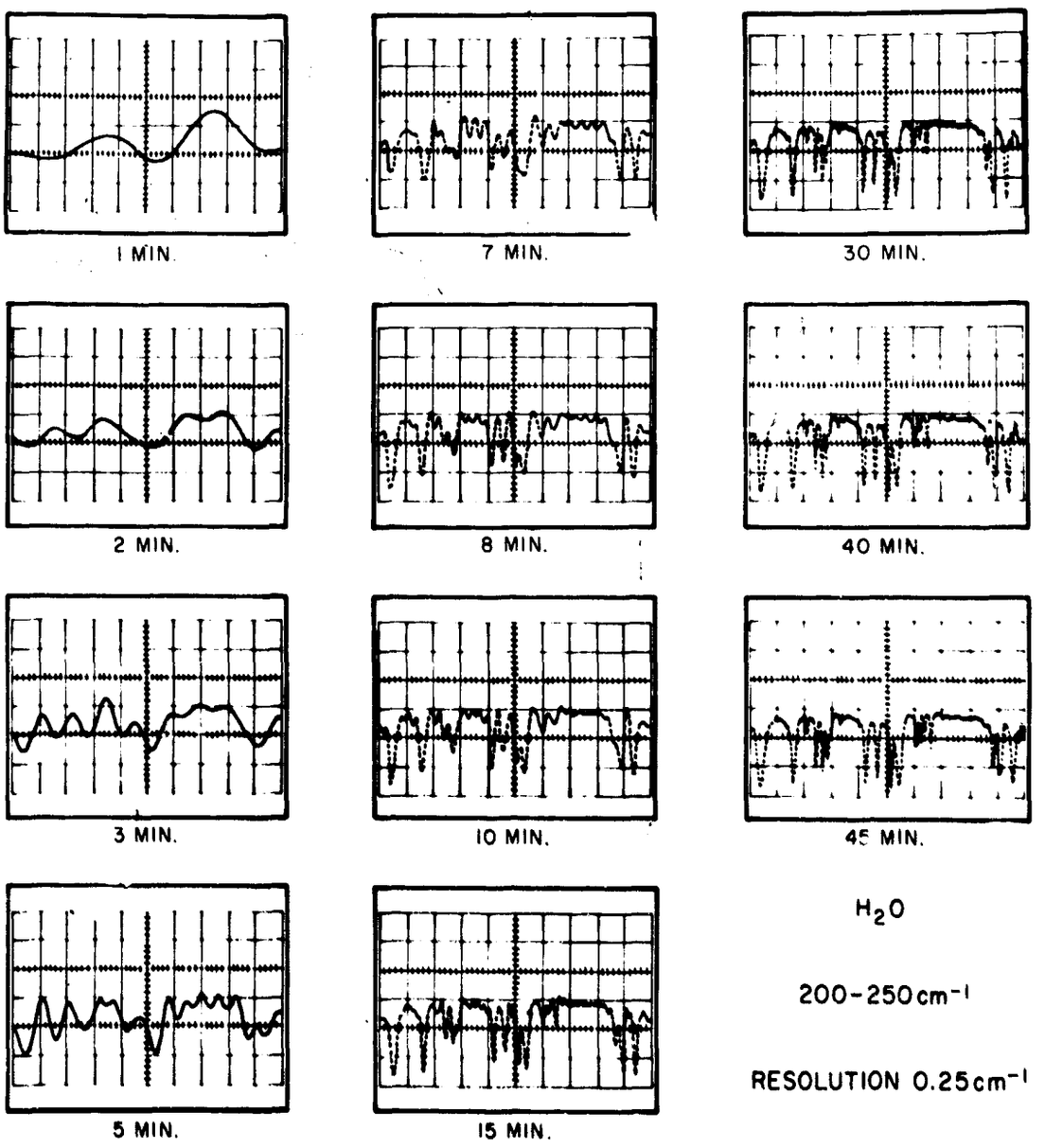


Figure 34-4. Photographs of Visual Display Screen Taken at Various Intervals During the Scan of a High Resolution Water Vapor Interferogram

where $p = (\Delta\tilde{\nu}\,\Delta x)^{-1}$ and is a constant for any particular Fourier transform computation, and i takes a range of integral values of $\nu_0\,\Delta\nu$, $\nu_0\,\Delta\nu+1, \ldots, \nu_0\,\Delta\nu+(M-1)$. For each new value of n the computer fabricates the range of values ni/p, determines the sign of the cosines and reduces the values of ni/p to the equivalent angle falling in the first quadrant, $0-\pi/2$.

The "reduced" values of ni/p are then compared with a prestocked table of 1024 ni/p values in the computer memory whose cosine values are arrayed between 1 and 0 at regular intervals of 2^{-10}. The computer finds the cosine of the tabulated ni/p value which gives the closest match to the fabricated value, by an incremental method of search.

In order to reduce the number of operations for fabricating a series of ni/p values at a given value of n, new values are quickly formed by the simple addition of $n\,\Delta i/p$ to the preceding value in the series, where Δi is the basic increment by which i changes.

(5) The apodized interferogram values

$$I'(n\,\Delta x)A_q(n/N)$$

are weighted by the appropriate value of cosine $(2\pi ni/p)$. The product is generated by a method of incremental addition which proceeds in step with the search routine of the table of cosine values, and obviates the need for the separate computer operation of multiplication. The process of computation is simpler and faster as a result.

(6) The values of the products formed in step 5 are stored in the separate addresses of the computer memory allocated for the range of spectrum points M. In the system described here, up to 1500 frequency points can be stored for a single beam spectrum. As successive new values are formed corresponding to additional samples of the interferogram, the data in each address are read out, modified by the addition of the new value and the accumulated result is rewritten into the same address.

(7) The spectrum is displayed on a cathode ray screen after each successive interferogram point, by means of a high speed D/A converter (see for example Figure 34-4).

(8) When the predetermined number of interferogram points have been scanned or when the accumulated spectrum is considered to have reached the desired quality as evidenced by the display screen, the array of spectrum values stored in the computer memory can be plotted on a strip chart recorder, by means of the D/A converter. A second pen on the strip chart recorder synchronized to the computer readout rate marks out the frequency scale at regular predetermined intervals (see for example, Figure 34-5).

(9) When a double-beam ratioed spectrum is to be obtained, up to 750 frequency points can be allocated in the computer memory for the "Background" and "Sample" spectra. The ratioed value of the two spectra is then obtained as a strip chart recording of an analog signal between 0 V and 5 V, which corresponds to ratio values between 0 and 100 percent.

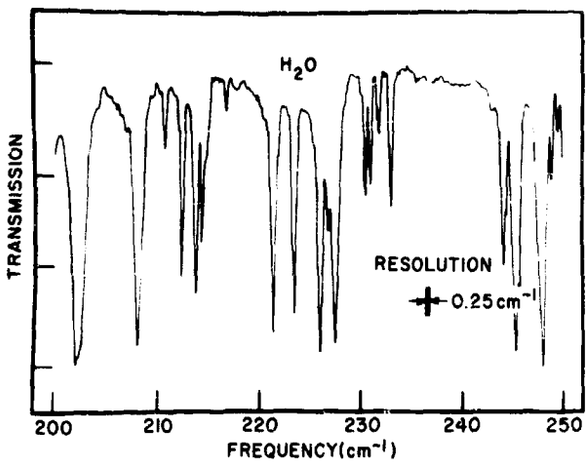


Figure 34-5. Chart Recording of Water Vapor Spectrum Obtained After 45 Minutes Shown in Figure 34-4

The computer memory locations of 4096 16-bit words are distributed as follows: frequency locations, 1500 words; table of cosine values, 1024 words; program and instructions, 1572 words.

The time of digital computation T is found to be almost directly proportional to the number of spectrum points calculated M, and follows an empirical relation

$$T \simeq 0.05 + 1.2\times10^{-3}\,M \text{ sec.} \qquad (34\text{-}5)$$

The time for calculation of 1000 frequency points is thus about 1.25 sec, as compared to 0.45 sec for the special purpose computer of Yoshinaga et al.[4] in which the 1000 frequency points are computed in parallel.

The rate at which spectra are plotted depends upon the intensity variations and the response time of the chart recorder, but it normally falls in the range 2 to 5 frequency points per second. The time to plot out a full spectrum of 1500 frequency points is thus about 5 to 12 minutes.

The limiting precisions of the various operations involved in the data logging and computing system may be listed as follows:

Quantization accuracy of A/D converter	±1 in 5000
Accuracy of estimating cosines from table	±1 in 2000
Stated accuracy of D/A converter	±1 in 2000
Accuracy of pen recorder	±1 in ~200

Since in practice, quantization errors of the A D converter and errors in determining the cosine terms may be assumed to be fairly random, the calculated accuracy of a summation involving a hundred or so terms becomes very much higher than that of a single term, which may be estimated at ±1 part in about 1800 from the above figures. It is obvious that the accuracy of calculation is considerably higher than the limiting accuracy of the pen recorder used to plot spectra.

34-4 RESULTS

It is instructive to discuss the performance of the interferometer-computer system in terms of the speed, accuracy, and versatility with which results can be obtained. Clearly, the maximum rate at which interferograms can be scanned and transformed in real time is limited first by the speed at which interferograms can be scanned without serious distortion, and second by the speed at which the desired array of frequencies can be calculated compared to the rate of arrival of successive interferogram data points. Table 34-1 lists typical speeds at which the interferometer can be used to scan different ranges of the far-infrared spectrum, together with the maximum interferogram sampling intervals that can be used in each case, to minimize the rate of data arrival.

Table 34-1 also shows the maximum number of frequency points that can be calculated under the various experimental scanning conditions, using the relationship expressed in Eq. (34-5).

It should be noted that the real-time method of calculation does not necessitate the storage of interferogram values in the computer memory, and there is thus no restriction on the number of interferogram values that can be processed. This allows the maximum resolution capability of the interferometer to be fully utilized. Furthermore, in cases where the number of spectrum points available in real time is unduly restricted by the rapidity of arrival of interferogram data points, the desired maximum 1500 points can be obtained by either reducing the interferogram scan speed or by employing optical filters to restrict the

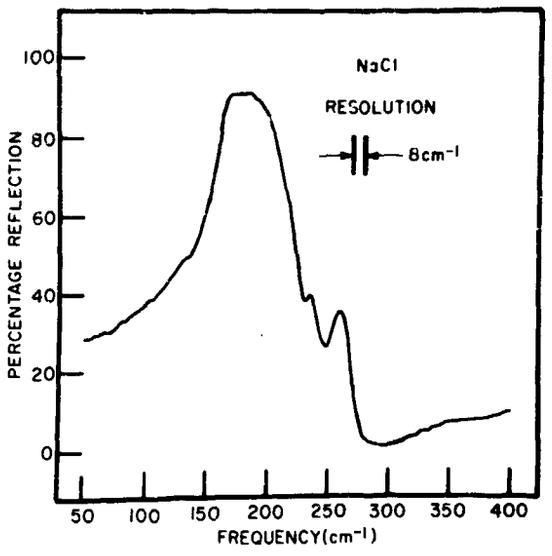


Figure 34-6. Double Beam Reflection Spectrum of an NaCl Single Crystal. (Scan time: 60 sec/interferogram)

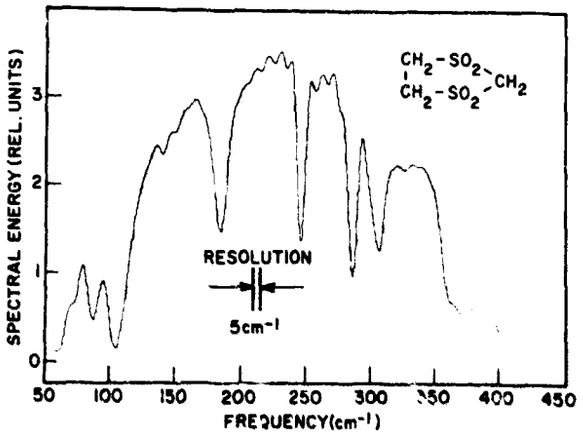


Figure 34-7. Single Beam Transmission Spectrum of a Powdered Sample of Cyclo Ethyl Di-sulphone Recorded in 3 Minutes

Table 34-1. Operating Parameters for Interferometer-Computer System

Frequency Range	Typical Interferogram Scan Speed	Sampling Interval	No. of Samples/Sec	Maximum No. of Frequency Points in Real Time
0–1000 cm^{-1}	2 μ/sec	5 μ	0.4	1500*
0–500 cm^{-1}	5 μ/sec	10 μ	0.5	1500*
0–250 cm^{-1}	10 μ/sec	10 μ	1.0	800
		20 μ	0.5	1500*
0–100 cm^{-1}	20 μ/sec	20 μ	1.0	800
		40 μ	0.5	1500*
0–50 cm^{-1}	50 μ/sec	40 μ	1.25	625
		80 μ	0.63	1300

*Maximum storage space allocated in memory

optical bandwidth of frequencies and using a larger interferogram sampling interval as allowed by information theory. Figures 34-5 to 7 show far infrared spectra recorded at various resolutions under typical scanning conditions, which illustrate the capabilities of the system for routine far-infrared spectroscopy.

The photometric accuracy of the interferometer real-time computer system described here is determined principally by the precision with which an interferogram sampling point can be arranged to coincide with the true zero-path interferogram maximum, since the present method of computation does not take account of phase errors.[8,9] For the cosine Fourier transform, such effects produce both an asymmetry in the instrumental scanning function and a distortion in the spectrum baseline, which becomes exceedingly pronounced at higher frequencies. To obtain less than 1 percent distortion of the highest frequency present ν_{max}, it is necessary for the phase error ϵ to be $\lesssim 1$ percent of the sampling interval, which is a difficult condition to satisfy in practice.

With the fine adjustment facility of the moiré fringe system used in this instrument, it is possible with a little care to record interferograms consistently with phase errors estimated to be less than 5 percent of the sampling interval, as judged by the sampling interval, as judged by the symmetry of values on each side of the steep zero path interferogram maximum. For the smallest sampling interval of 5 μ, phase errors of this magnitude could lead to intensity distortions about 10 percent full scale or more at 1000 cm^{-1} but considerably less at frequencies below 500 cm^{-1}.

To illustrate the precision with which interferograms can be recorded, Figure 34-8 shows a high

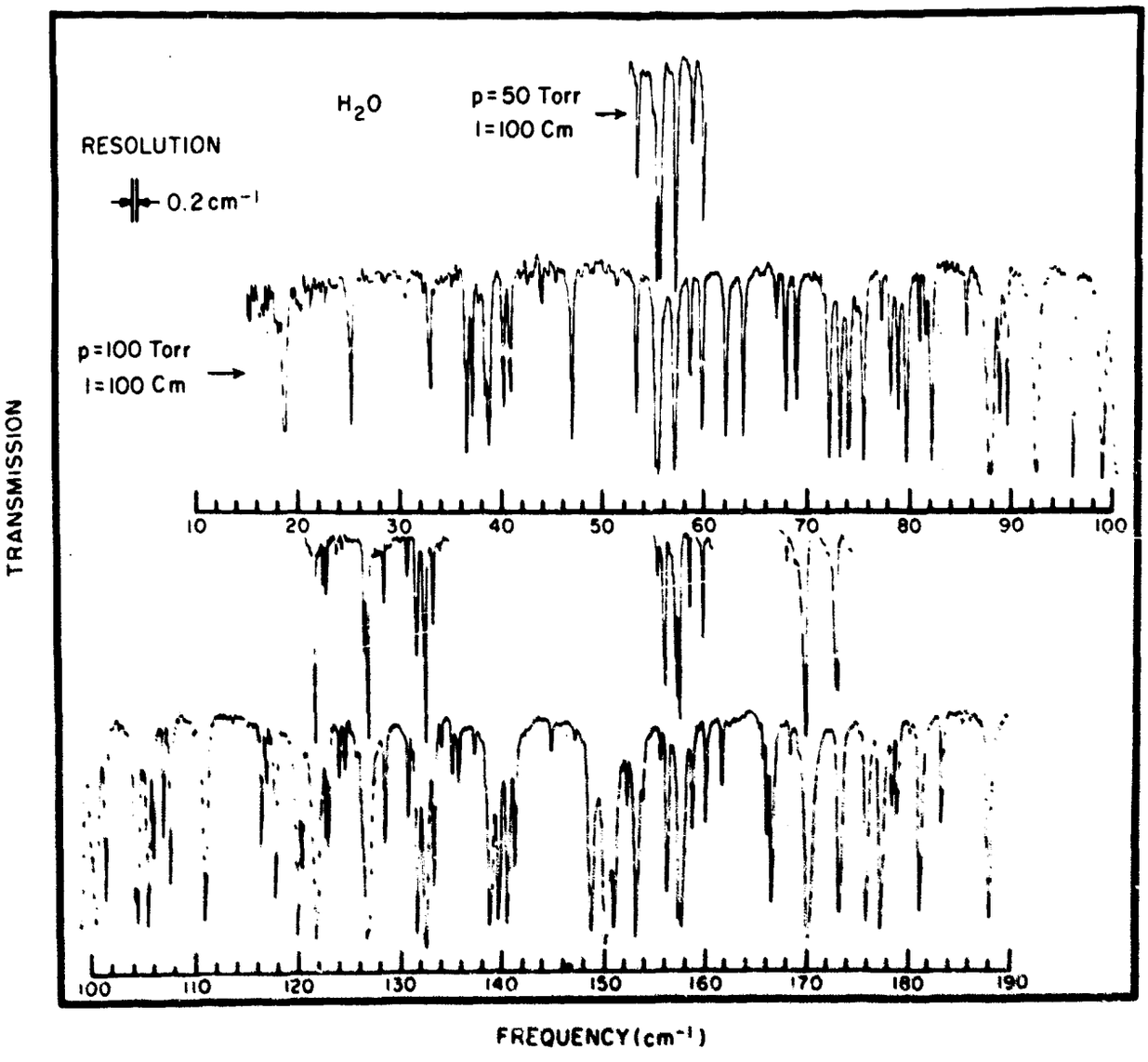


Figure 34-8. High Resolution Water Vapor Spectrum, Assembled From Three Separate Double-Beam Scans. (Total recording time 3 hours)

resolution double beam water vapor spectrum obtained as a composite of spectra recorded over three separate frequency ranges. All the well documented[10] water vapor absorption lines are reproduced with the expected relative intensities, and the spectrum shows little or no effects of distortion as evidenced by the symmetry of the lines and the flatness of the transmission baseline. Figure 34-9 shows a high frequency mercury arc spectrum recorded in the 100- to 950-cm^{-1} spectral region. This also shows little evidence of phase errors as the strong absorption bands of polyethylene and Melinex in the 700- to 900-cm^{-1} region flatten near zero and do not take negative values as would normally be the case for an incorrectly sampled interferogram.

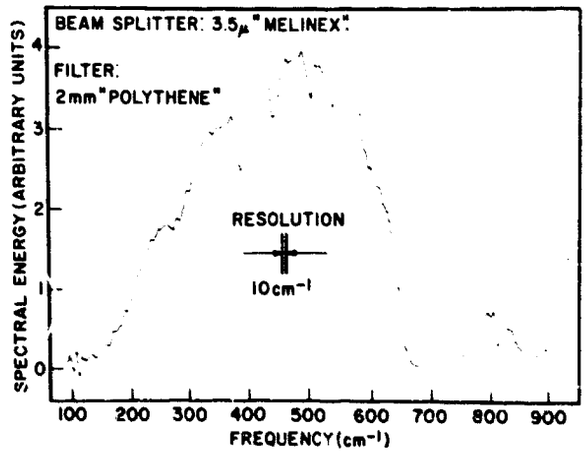


Figure 34-9. High-Frequency Single-Beam Instrumental Background Spectrum Recorded in 10 Minutes

Other limitations to the photometric accuracy of the system described here are common to other interferometric spectrometers and arise mainly from effects of source or detector drift during and between the scanning of interferograms. Such drift causes the 100-percent level for double-beam transmittance or reflectance measurements to be in error, and may also introduce noticeable sinusoidal distortions in the computed spectrum due to the shift of the interferogram baseline away from zero volts. These effects can be minimized by water cooling the Golay detector and by running the source in a permanent vacuum. With care, such intensity errors may be restricted to one to two percent full scale.

34-5 CONCLUSIONS

The 4K memory general purpose digital computer system described here has proved fully adequate for the needs of routine far-infrared Fourier spectroscopy by virtue of its full digital accuracy and the large number of spectrum points that can be calculated at one time. Furthermore, its ability to be programmed in a real-time mode of operation reduces the spectrum acquisition time to a minimum, and gives the operator full control to evaluate and determine the course of his experiment. The principal limitation of the real-time method of programming described here is the lack of a phase error correction procedure, which could allow noticeable intensity errors to be introduced particularly at frequencies $>500\ cm^{-1}$ unless the interferometer is very carefully adjusted for the exact coincidence of a sampling point at the zero path position. It seems unlikely at this moment, however, that phase error correction procedures can be carried out without endangering the speed of the real-time method of computation, or substantially reducing the number of frequency points that can be calculated.

References

1. Strong, J., and Vanasse, G. A. (1958) *J. Phys. Radium* **19**: 192.

2. Pritchard, J. E., Bullard, A., Vanasse, G. A., and Sakai, H. (1967) *J. de Physique* **28**, C2: 67.

3. Edgar, R. F., Lawrensen, B., and Ring, J. (1967) *J. de Physique* **28**, C2: 73.

4. Ridyard, J. N. A. (1967) *J. de Physique* **28**, C2: 62.

5. Yoshinaga, H., Fujita, S., Minami, S., Suemoto, Y., Inoue, M., Chiba, K., Nakano, K., Yoshida, S., and Sugimori, H. (1966) *Appl. Opt.* **5**: 1159.

6. Hoffman, J. E. (1969) *Appl. Opt.* **8**: 323.

7. Chantry, G. W., Evans, H. M., Fleming, J. W., and Gebbie, H. A. (1969) *Infrared Phys.* **9**: 31.

8. Connes, J. (1961) *Rev. Opt.* **40**: 72.

9. Loewenstein, E. V. (1963) *Appl. Opt.* **2**: 491.

10. Hall, R. T., and Dowling, J. M. (1967) *J. Chem. Phys.* **47**: 2454.

Discussion

Q. (L. W. Thorpe, Beckman R.I.I.C.): What are the relative sizes of core store used in the Varian computer for the program, data storage, and instrument control, and so forth?

A. (R. Milward): The memory locations of the 4K Varian 620/i computer used for real-time Fourier spectroscopy are allocated as follows:

Frequency locations	1500 words
Table of cosine values	1024 words
Program and instructions	1572 words

where all words are of 16-bit magnitude.

Q. (L. Mertz, Smithsonian Astrophysical Observatory): Does phase information have to be available *a priori*?

A. (R. Milward): The real-time method of transformation described here does not make any provision for either the estimation or correction of phase errors, and demands the exact coincidence of a sampling point with the zero path interferogram summit.

Contents

35. Application of Phase Correction to Multiplexing Fourier Spectroscopy

L. W. Thorpe
Beckman R.I.I.C. Ltd
Worsley Bridge Road
London, S.E. 26, England

Abstract

A method for digital correction of phase errors existing in the interferogram obtained from a multiplexing Fourier spectrometer is discussed. The technique demonstrated uses interferograms obtained from a mock interferometer. The whole process is shown, from the input interferogram to the complete corrected interferogram, with the effect on the spectrum. There is a discussion of a method of transforming a large data array in only half the computer memory normally required.

35–1 INTRODUCTION

Fourier spectroscopy is the technique of recovering the power spectrum of a particular radiative source by Fourier transformation of a recorded interferogram. The most commonly used Fourier spectrometers are both interferometric and multiplexing, and include the Michelson and lamellar grating interferometers. Most of the interferograms discussed here were obtained using a mock interferometer which is multiplexing and Fourier but not interferometric. The term Multiplexing Spectrometer will be used to refer to the three instruments upon which the computations process has been proved.

The Fourier transform technique has received considerable application in the last decade in the fields of far-infrared spectroscopy and infrared astronomy, where the multiplex advantage[1] helps to overcome the difficulties encountered in the measurement of weak radiative sources. The principal disadvantage of the method is that the spectra of interest are not directly available, but have to be recovered from lengthy Fourier transform computations performed on a digital computer.

The earlier Fourier transform programs used for this purpose were often time consuming and therefore costly and, furthermore, they did not make adequate provisions for correction of phase errors and other instrumental errors incurred during the recording of the interferograms. These instrumental errors and their effect on the computed spectra have since been studied in detail by J. Connes.[2]

Recently, fast Fourier transform routines which reduce the computation time for transforms involving many data points, as well as more powerful methods of correcting interferograms have been devised.[3–6]

The use of a correction technique involves a fairly large computer, and the necessity of efficient storage of the program and data becomes acute. This paper deals with a general-purpose fast program to correct the errors described above, and with the development of a routine to transform a very large number of data

points when only a portion of the output spectrum is required.

35-2 CORRECTION TECHNIQUE PROGRAM

In the ideal case the interferogram produced from the types of multiplexing spectrometer discussed above should be a perfectly symmetric function. In practice however, the digitally recorded interferogram may be asymmetric due to the imperfect sampling of the interferogram at and around the zero-path difference position, while real asymmetry can be produced when a dispersive medium is placed in one arm of the Michelson interferometer, which causes the optical path differences to vary with the frequency of the radiation.[7,12] In the ideal case, the spectral distribution is given by the cosine Fourier transform of the interferogram from the zero-path difference out to the final sample point. In the nonideal case the spectral distribution is usually obtained from the asymmetric recording as the modulus of the complex Fourier transform, using an interferogram with an equal number of points on either side of the zero-path difference position.

This process was shown by Connes[2] to have two disadvantages: (a) the interferometer must be capable of producing twice the travel and twice the number of sample points required in the ideal case; (b) the production of the modulus of a complex number is a nonlinear process and the noise in the final spectrum is no longer constant, but increases as the signal decreases. It was shown by Forman, Steel, and Vanasse,[3] that an asymmetric interferogram could be corrected if it is assumed that a function $\phi(V)$ exists and represents the phase errors present in either the equipment or the associated recording electronics or both. The phase function $\phi(V)$ is calculated from a portion of the interferogram, beginning at a small displacement before the zero-path difference position to an equivalent point after the zero-path difference position.

From the phase function the correction function is obtained by Fourier transformation of $\exp[i\phi(V)]$. This function must be apodized and convolved across the input interferogram in order to form an autocorrelation function with little or no asymmetry. The apodization of the correction function is essential as this reduces the modulation at the ends of the correction function, which otherwise produces spikes in the corrected interferogram and hence spurious ripples in the final output spectrum.

To find the optimum apodization function to be used with the Fourier transform of $\exp[i\phi(V)]$, a number of tests were performed using analytic data which had known phase errors present. The analytic interferogram was developed by the computer to represent a bandpass box-car function similar to the bandpass of a mock interferometer, with two absorption lines and an emission line. This type of test spectrum was selected as it compares with the band pass spectrum of a mock interferometer when used in the study of emission lines present on a continuum containing atmospheric absorption lines. It can also be used to assess the effects of the computational procedure upon cut-on and cutoff regions of a spectrum and the effect upon the ratio of known level in the spectrum. Figure 35-1A shows this test function when the input interferogram had no phase errors present, and this can be used as a comparison with the output from the correction procedure when the input interferogram contained phase errors linearly dependent upon frequency, and the correction function has been apodized with the specified function. Figure 35-1B shows the cosine Fourier transform of the input interferogram when it contained phase errors linearly dependent upon frequency.

Figure 35-1C shows the transform of the corrected interferogram when the correction function was apodized with an $\exp(-X^2)$ function, and Figure 35-1D shows the transform of the corrected interferogram when the correction function was apodized with a $(1-X^2)^2$ type of function.

This correction technique is only meaningful when the signal-to-noise ratio in the portion of the interferogram around zero-path is adequate (>30 say) to produce a smooth phase function in the region of frequency space involved in the input data. If the signal-to-noise ratio is not sufficient and the multiplexing spectrometer is stable with time, it is possible to precalibrate the phase function, and hence the correction function, using a high-energy radiative source. The correction function obtained in this way may then be used to correct low signal-to-noise ratio interferograms without producing nonlinear effects on the signal-to-noise ratio in the spectrum.

The Cooley-Tukey algorithm[8] was used during the tests and also in the development of a general-purpose computer program for the correction of asymmetric interferograms obtained in Fourier spectroscopy. The complete program has been used with considerable success to correct interferograms obtained from the three multiplexing spectrometers mentioned earlier, and results are shown of the mock interferometer and Michelson interferometer outputs.

The mock interferometer[9-11] has been used in the laboratory on known emission line sources and has been subject to preliminary tests in the field of observational astronomy. The results gave instrument calibration spectra and "blue sky" spectra in the near infrared. The instrument is to be used in the near future to obtain spectra of gaseous nebulae in the near infrared and to investigate the transmission properties of the earth's atmosphere.

Figure 35-2 shows the complete correction process performed on an interferogram obtained from a laboratory line source, and a discussion of its layout follows. The graphs at the upper left and the lower left show the complete input interferogram, and the region around the zero-path difference position used in the phase error computations, respectively. It can be seen from these two graphs that this interferogram is antisymmetric and has considerable low-frequency drift in the region around the zero-path difference position. The graph at the lower center shows the

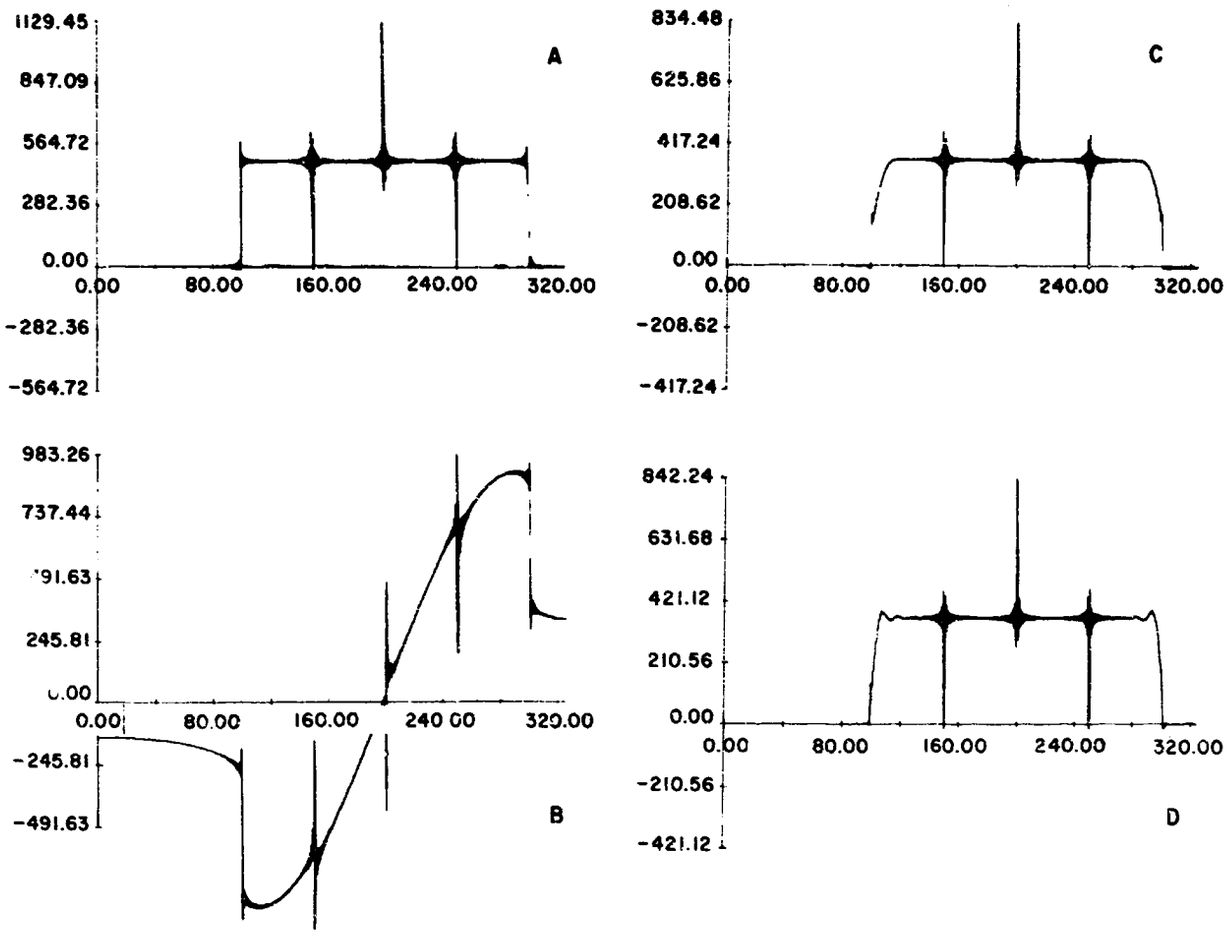
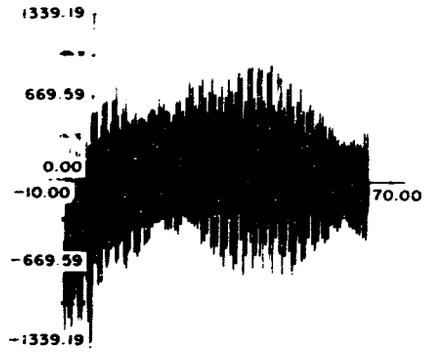


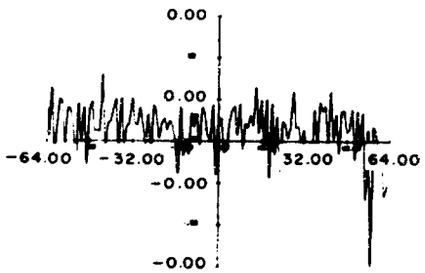
Figure 35-1. Transforms of Complete Corrected Interferograms After Using the Normal Distribution and the Connes Apodization Function on the Cosine Correction Function Compared With the Transform of the Uncorrected Interferogram, and with an Interferogram Containing No Phase Errors

phase function computed from the arctangent of the ratio of the sine and cosine transform of the region around the zero-path difference position, shown in the graph at the lower left. These interferogram points are transformed by the Cooley-Tukey algorithm with the zero-path difference member located in the first point of the cosine array and the points to the right located in the next sixty-three points of this cosine array. The points to the left of the zero path difference position are located in the remaining sixty-four locations of the cosine array such that the point on the immediate left of the zero-path position is in the one hundred and twenty-eighth location of the cosine array. The complete sine array at the input to the Cooley-Tukey routine is put to zero, and a binary order seven transform is performed. The output arrays from the Cooley-Tukey algorithm contain the zero frequency components in their first locations with the positive frequency components out to the alias or sixty-fourth output point. The remaining sixty-four points of each array represent the amplitudes of the negative frequency components and are required for this computation. It is found that the cosine array is symmetric about the zero-frequency components while the sine array is antisymmetric about these components. Consequently, the phase function is antisymmetric as required in the analysis produced by Forman, Steel, and Vanasse.[3]

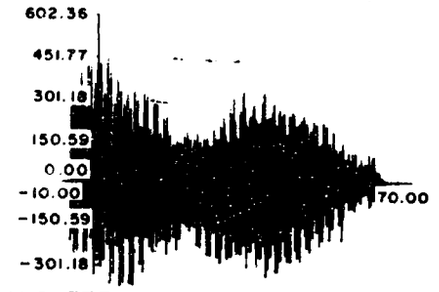
The center-center graphs of Figure 35-2 show the apodized correction function obtained from the cosine array at the output of the transform of exp $[i\phi(V)]$. The components of exp $[i\phi(V)]$ prior to its Fourier transform are entered into their respective positions of the cosine and sine arrays with the zero and positive frequency components in the first sixty-four points and the negative frequency components in the last sixty-four points of each array. As would be expected with two input arrays of this form, that is, the cosine array symmetric about the zero frequency component and the sine array antisymmetric about the zero frequency component, the transformed arrays will only contain meaningful data in the cosine transform and the sine transform will be zero. This is shown to be the case by the sine correction function graphed in the upper center position of Figure 35-2, which shows all zero's on the Y axis.



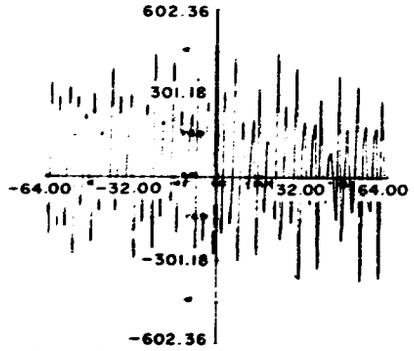
COMPLETE INPUT INTERFEROGRAM



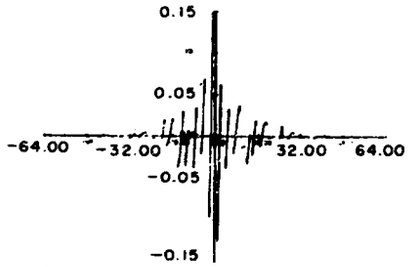
SINE CORRECTION FUNCTION



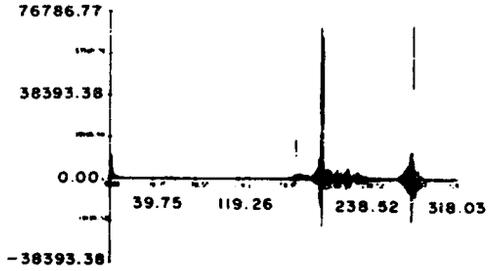
COMPLETE CORRECTED INTERFEROGRAM



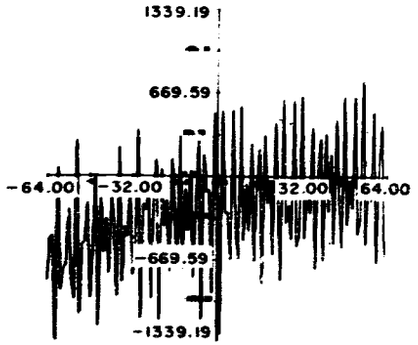
CORRECTED INTERFEROGRAM



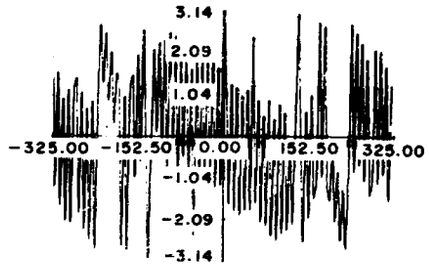
COSINE CORRECTION FUNCTION



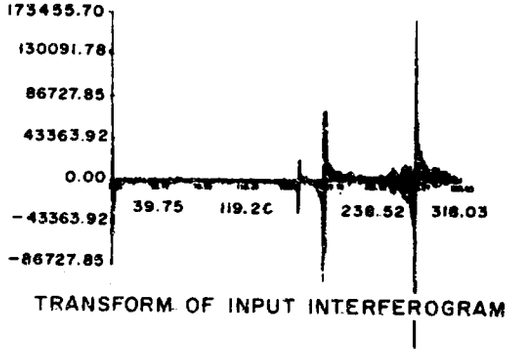
TRANSFORM OF CORRECTED INTERFEROGRAM



INPUT INTERFEROGRAM



INPUT PHASE FUNCTION



TRANSFORM OF INPUT INTERFEROGRAM

Figure 35-2. Phase Error Correction of Laboratory Interferogram NAY 8321 Obtained From a Mock Interferometer and a Laboratory Line Emission Source From 0.4 to 0.75 Microns

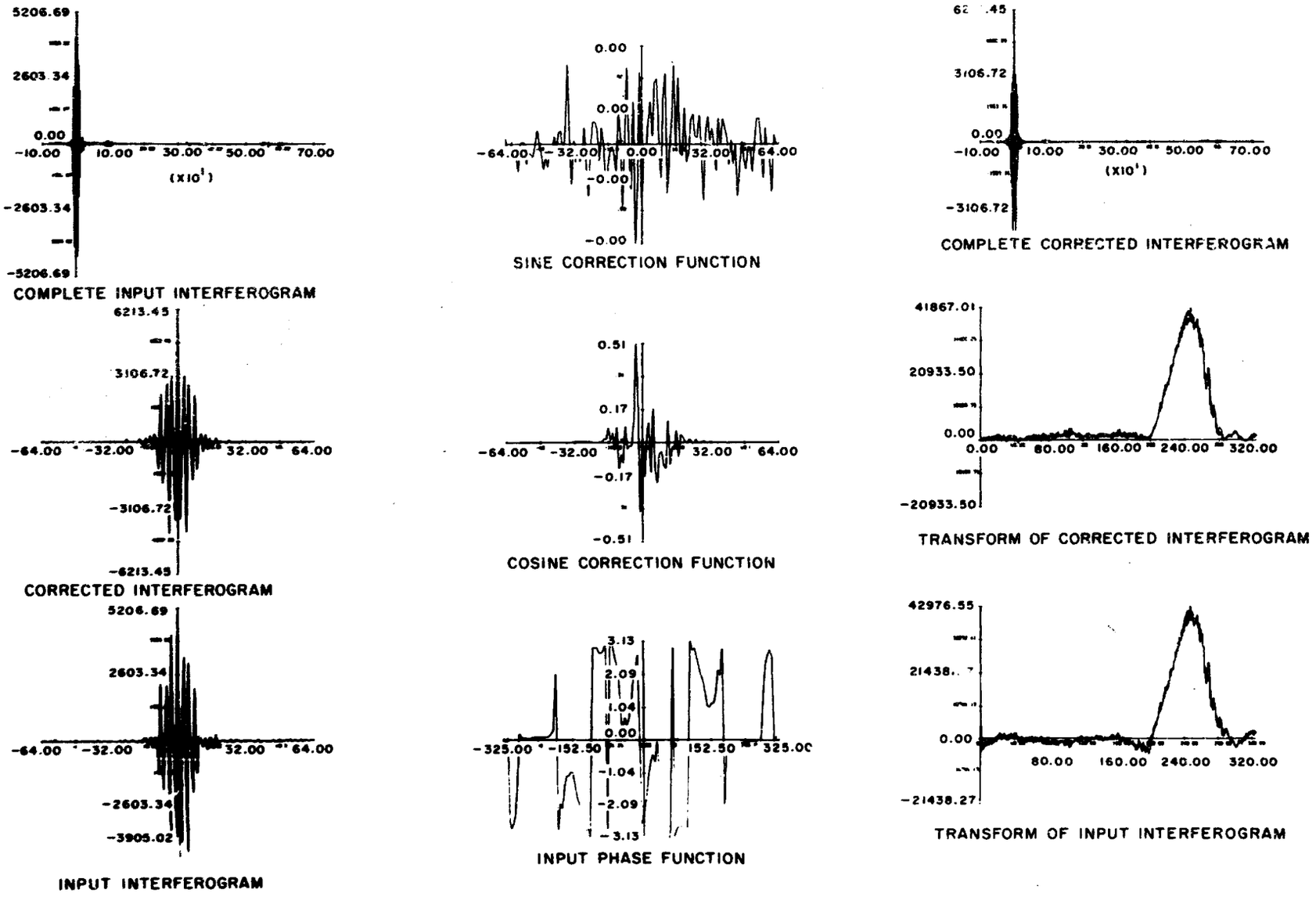
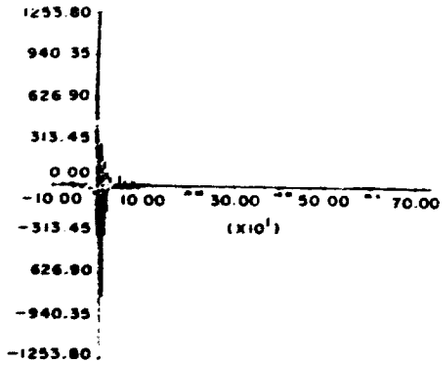
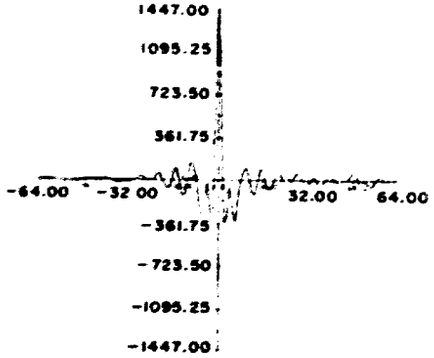


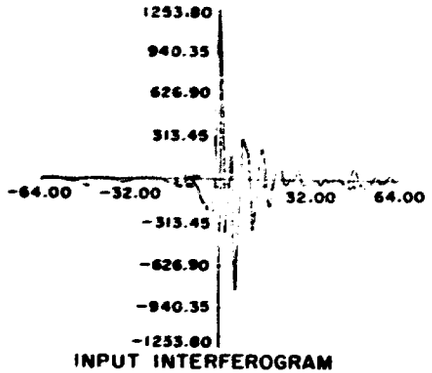
Figure 35-3. Phase Error Correction of Sil. Solar (3) 1.00 Interferograms Obtained From a Mock Interferometer and the Blue Sky Radiation From 0.8 to 1.0 Microns



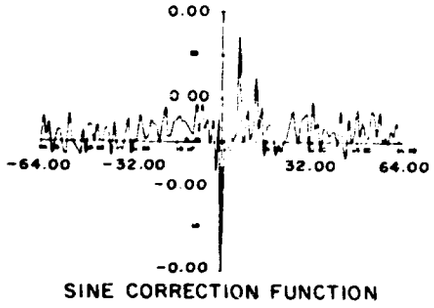
COMPLETE INPUT INTERFEROGRAM



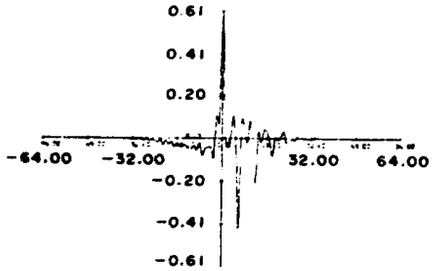
CORRECTED INTERFEROGRAM



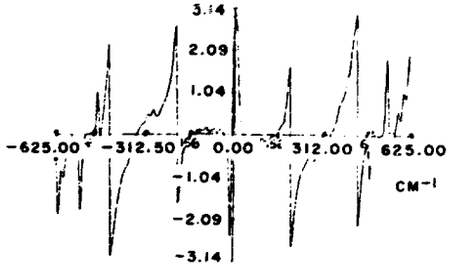
INPUT INTERFEROGRAM



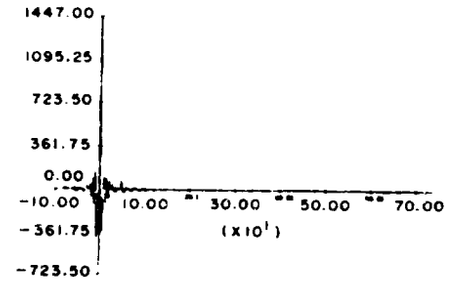
SINE CORRECTION FUNCTION



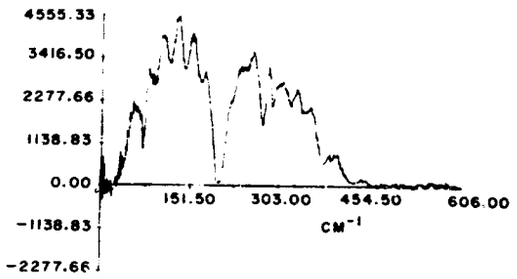
COSINE CORRECTION FUNCTION



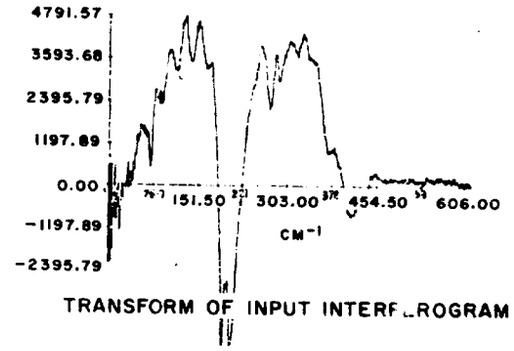
INPUT PHASE FUNCTION



COMPLETE CORRECTED INTERFEROGRAM



TRANSFORM OF CORRECTED INTERFEROGRAM



TRANSFORM OF INPUT INTERFEROGRAM

Figure 35-4. Phase Error Correction of Asymmetric Teflon Interferogram Obtained From a Michelson Interferometer Using 0.150 mm Thick Teflon Inserted in One Arm of the Interferometer in the 30 to 00 (cm^{-1}) Region

The cosine correction function shown in the center-center position is then convolved across the input interferogram to produce the symmetrized, corrected interferogram shown in the top right and center left graphs of Figure 35-2. The cosine Fourier transforms of the input interferogram and the corrected interferogram are compared in the lower right and center right graphs of Figure 35-2 and demonstrate how effective the process is even on line spectra, and that the signal-to-noise in the transform of the corrected interferogram appears better than in the transform of the uncorrected interferogram. The results obtained are at the expected resolution and the instrumental profile can be seen to be symmetrical. The transformation of both corrected and uncorrected interferograms is performed with the zero path difference element located in the first member of the cosine array, with the point before zero-path difference ignored.

Figure 35-3 demonstrates the complete correction process performed on a second mock interferogram, and shows that features on the sides of the pass band are not grossly distorted in magnitude or shape by the correction process.

Figure 35-4 demonstrates the correction process used with an interferogram obtained with a Michelson interferometer. An asymmetric interferogram was deliberately produced by placing the sample in one arm of the interferometer and thus producing path differences dependent upon the change of refractive index of the material with wavelength. The phase function and the transform of the corrected and uncorrected interferograms have their X axes calibrated in wavenumbers, the alias frequency being 625 cm^{-1}, corresponding to an 8-μ sampling interval.

It can be seen that the discontinuity in the phase function from 1.5 to -3.0 rad is at the 200 cm^{-1} absorption band of Teflon and that the phase function can be used for accurate refractive index determination.[7,12] The phase correction of these three different types of interferograms demonstrates the usage capability of the technique described in Forman, Steel, and Vanasse.[3] The procedure has given symmetrical interferograms in all cases; it has helped to reduce the noise amplitude in the region of very small or zero signal levels; it has removed the effect of small long-term drift in the interferogram; and it has given a method for obtaining from one interferogram both the refractive index changes of a sample with wavelength, and its absorption spectrum. It is interesting to note that for each of the cases discussed, the cosine correction function bears a slight similarity to the portion of the interferogram around the zero-path difference region, but it is in no case a direct copy of this region of the interferogram.

For the correction of interferograms obtained using either the mock or the Michelson interferometer the number of points used in the correction function becomes important due to two effects. The first is that, since a convolution takes place, the larger the number of points in the correction function the larger the number of operations in the convolution process and hence the larger the computation time. The second effect is the broadening and distortion on absorption lines produced by having too high a resolution in the phase function. If the resolution is too high, the phase function is no longer a smooth function dependent upon the interferometer alone, but becomes distorted because the data in the region around the zero-path difference position actually start to resolve the absorption line out of the continuum, and hence change the phase spectrum.

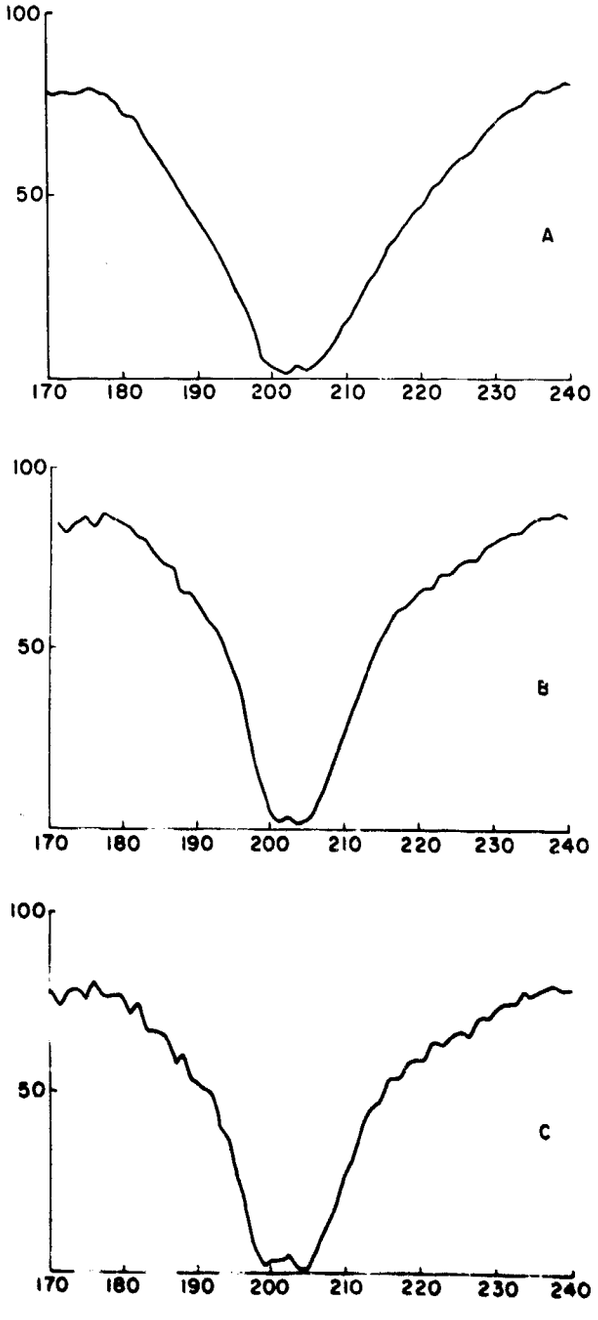


Figure 35-5. Effect of Different Numbers of Points in the Phase Function

The correction function will contain information about the changes in the phase function and the effects will be noticed in the transform of the corrected interferogram. The effect of different numbers of points in the phase function is shown in Figure 35-5, which compares A, the width of the Teflon absorption line obtained when 128 points are used in the correction function, with B, the width obtained when 32 points are used. These are then compared with the width obtained using no correction technique, but producing the power transform shown as Figure 35-5C. It can be seen that the widths shown in Figures 35-5B and 35-5C are the same, while that shown in Figure 35-5A is larger. It is also noticeable that the noise content shown in Figures 35-5A and 35-5B is less than that shown in Figure 35-5C.

It is therefore necessary to decide what is the width of the widest feature required in the spectrum, and what is the width of the narrowest feature required in the spectrum. From the first width, the number of points in the correction function can be calculated, and in general will be approximately half the number required to resolve this feature. The second width defines the number of points in the complete interferogram and this is related to the final resolution required in the output spectrum.

35–3 PARTITION AND CONJUGATION

This section deals with the transformation of input data points that have been corrected for any phase errors, and from which only a portion of the output spectrum is required. Consider a function $f(n\,\Delta x)$ to represent the array of points to be analyzed. Let the odd numbered members of the array be inserted into the input real array, with alternating sign, and let the even numbered members be inserted into the input imaginary array, with alternating sign. This process can be represented by

$$G(n\,\Delta x) = f(n\,\Delta x)\exp(i\pi n/2) \tag{35-1}$$

for $0 \leq n \leq N = 2^{(M-1)}$. Consider the number of data points being transformed to be 1024, which for the normal fast Fourier transform method would require $M = 11$ and the data points to be inserted into the first 1024 points of the input real array. For the method being discussed, $M = 10$ and the odd numbered members of the array $f(n\,\Delta x)$ are inserted into the input real array while the even numbered members are inserted into the input imaginary array.

Since the function $G(n\,\Delta x)$ has real and imaginary parts, its transform will be complex, and the four output arrays corresponding to the positive and negative frequency components of the real and imaginary output arrays are related to the cosine and sine transforms of the two functions in $G(n\,\Delta x)$ by the following equations

$$G(n\,\Delta x) = A(n\,\Delta x) + iB(n\,\Delta x). \tag{35-2}$$

$$\begin{aligned} g(\nu) &= \int_{-\infty}^{+\infty} G(n\,\Delta x)e^{-2\pi i\nu n\,\Delta x}\,dx \\ &= \int_{-\infty}^{+\infty} A(n\,\Delta x)e^{-2\pi i\nu n\,\Delta x}\,dx \\ &\quad + i\int_{-\infty}^{+\infty} B(n\,\Delta x)e^{-2\pi i\nu n\,\Delta x}\,dx \\ &= \{a_r(\nu) - b_i(\nu)\} + i\{a_i(\nu) + b_r(\nu)\} \\ &= g_r(\nu) + ig_i(\nu). \end{aligned} \tag{35-3}$$

For the negative frequencies

$$\begin{aligned} g(-\nu) &= \{a_r(-\nu) + ia_i(-\nu)\} + i\{b_r(-\nu) + ib_i(-\nu)\} \\ &= \{a_r(\nu) + b_i(\nu)\} + i\{b_r(\nu) - a_i(\nu)\} \\ &= g_r(-\nu) + ig_i(-\nu), \end{aligned} \tag{35-4}$$

from which the individual transforms are

$$a_r(\nu) = [g_r(\nu) + g_r(-\nu)]/2, \tag{35-5}$$

$$a_i(\nu) = [g_i(\nu) - g_i(-\nu)]/2, \tag{35-6}$$

$$b_r(\nu) = [g_i(\nu) + g_i(-\nu)]/2, \tag{35-7}$$

$$b_i(\nu) = [g_r(-\nu) - g_r(\nu)]/2. \tag{35-8}$$

Since alternate members of the input array have been inserted into the real and imaginary parts of G, the sampling interval has effectively been doubled and thus the alias frequency for the output spectra will be half that corresponding to the actual sampling interval. If, however, two unrelated functions be used as A and B, with the normal sampling interval, it

is possible to Fourier transform both functions simultaneously and to separate the cosine and the sine transforms of each function by the relations given above, to the alias frequency determined by the actual sampling interval. For the process under discussion, the effect of inserting the input array into the transform routine with the signs of alternate members changed, is to produce the output spectrum with the negative frequency components in the first half of the output arrays, and the positive frequency components in the second half of the output array for both the cosine and sine transforms.

From the equations above, if the addition of corresponding elements about the central member of the real output array is performed, the cosine Fourier transform of the input array is obtained to half the old alias frequency, or to 256 cycles in the input record, and the array will have 512 members. Similarly, if subtraction of corresponding elements about the central member of the imaginary output array is performed, the sine Fourier transform of the input array is obtained, again to half the old alias frequency, and the function will also contain 512 members. Using this method, the complex Fourier transform from zero frequency to half the old alias frequency at the same resolution as before is produced, but only half the previous amount of computer store and somewhat less computer time has been used. In this way it becomes possible to use a digital computer to transform twice as many data points as was previously possible; or similarly, the small computer may be programmed to process more data points in one program than was previously possible.

This technique was tested on a practical interferogram obtained from a lamellar grating interferometer with water vapor as the sample. Figure 35-6 shows the spectrum obtained using the normal transform procedure with $M = 11$, and the spectrum obtained of the lower frequency region, using $M = 10$. Figure 35-6 shows that the process produces satisfactory results. This technique offers the multiplexing high-resolution spectroscopist an increase in resolution of a factor of two, if his resolution is limited solely by the size of computer storage available and if the Fourier transform program is based on the Cooley-Tukey algorithm.

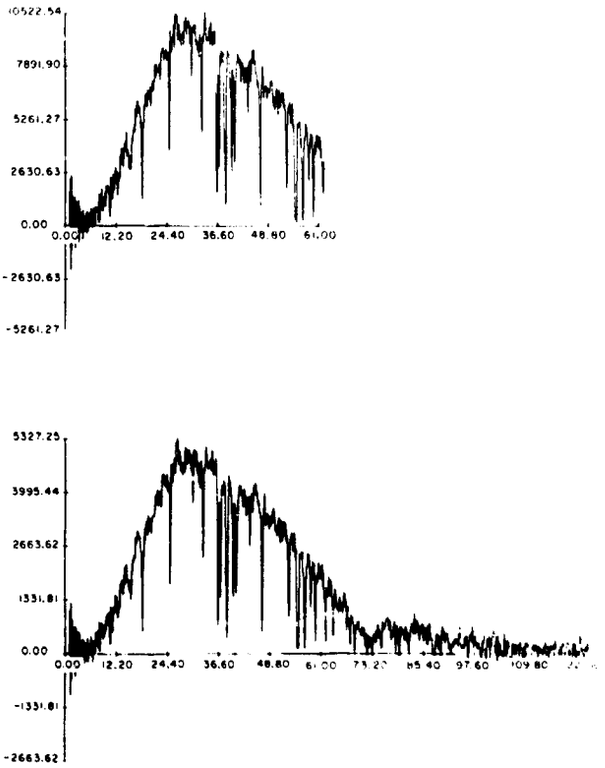


Figure 35-6. Transformation of a Lamellar Grating Interferogram of a Water Vapor Sample Using the Method of Partition and Conjugation With the Interferogram Prefiltered to Remove All Frequencies Above the New Alias Frequency

35-4 SUMMARY

A general discussion of the correction of phase errors occurring in the interferograms obtained from different forms of multiplexing spectrometers is developed, together with the effect of different apodization functions on the cosine correction function. A discussion is given of the effects on the transform of different numbers of points taken in the correction function, and also of a new use of the algorithm, whereby large arrays can be transformed to obtain high-resolution spectra of a selected region of the frequencies present in the input interferogram.

Acknowledgments

Part of the work discussed in this paper was developed in support of a Ph.D. thesis, and the author would like to thank the staff of the Department of Computation of Imperial College, London, for the use and advice on the programming of their IBM 7094 computer, without which the work would not have been possible.

The author wishes to thank Dr. M. J. Selby, Dr. D. J. Neale, and Dr. R. C. Milward for their assistance, valuable comments, and advice during this work; and Professor Ring and Dr. J. F. Gainger for their assistance during the preparation of the manuscript.

References

1. Felgett, P. (1958) *J. Phys. Radium* **187**: 237.

2. Connes, J. (1961) *Rev. Opt.* **440**: 45, 116, 171, 231.

3. Forman, M. L., Steel, W. H., and Vanasse, G. A. (1966) *J. Opt. Soc. Am* **56**: 59.

4. Sakai, H., Vanasse, G. A., and Forman, M. L. (1968) *J. Opt. Soc. Am.* **58**: 84.

5. Mertz, L. (1967) *Infrared Phys.* **7**: 17.

6. Bingham, C., Godfrey, M. D., and Tukey, J. W., Of audio and electro acoustics, *IEEE Trans.* **AV-15**: 56.

7. Chamberlain, J., Gibbs, J. E., and Gebbie, H. A. (1969) *Infrared Phys.* **9**: 185.

8. Cooley, J. W., and Tukey, J. W. (1965) *Math. of Computation* **19**: 297.

9. Mertz, L., *Transformations in Optics*, John Wiley and Sons Inc., London.

10. Thorpe, L. W. (1970) Thesis, Hull, England.

11. Selby, M. J., and Ring, J. (1966) *Infrared Phys.* **6**: 33.

12. Chamberlain, J. E., and Gebbie, H. A. (1966) *Appl. Opt.* **5**: 3.

Discussion

Q. (James L. Lauer): How many sample points do you typically use to determine the correction function you described for phase correction, when you are sampling at 8-μ intervals, to get optimum band shapes?

A. (L. Thorpe): As described at the end of Section 35-2, the number of points used in the correction function is dependent upon two factors. The first is that the larger the number of points in the correction function, the longer the computation time for the

convolution of the correction function across the input interferogram. The second is that, should the interferogram points of the input interferogram used in the calculation of the correction function start to resolve any broad bands of the spectrum, then distortions will occur on these bands as shown in Figure 35-5.

It is therefore important to use as small a number of points in the correction function as possible, and, for example, since the approximate width of the P.T.F.E. absorption band at 200 reciprocal cm is 20 reciprocal cm, the number of points in the correction function will be

$$N=\frac{1}{2}\frac{L}{\Delta X}=\frac{1}{2}\frac{1}{R\Delta X}=\frac{10^4}{2\times 20\times 8}\approx 32.$$

In general the number of points used in the correction function, for the work described in this paper, has been between ± 16 and ± 64 points about the zero-path difference position of the input interferogram, and has been calculated as above to not resolve the broadest feature required in the output spectrum.

Comment (L. Mertz): It is not fortuitous that the cosine correction function resembles the center of the interferogram. The only difference is the effective filter bandpass. In fact one can avoid computing the cosine correction function and simply use the central section of the interferogram. It is only necessary to then divide the resulting spectrum by the spectrum of that small central interferogram to restore the levels to the proper values.

Comment (L. Thorpe): The shape of the cosine correction functions shown in Figures 35-2 and 35-3 are quite dissimilar to the zero-path difference region of the input interferogram, while that shown in Figure 35-4 is somewhat similar. However, to convolve the input interferogram with its own zero path difference region and then divide the transform of this "autocorrelated" interferogram with the transform of the zero-path difference region produces a number of pitfalls.

(a) Either an interpolation of the output frequency points of the low order transform of the zero-path difference region must be performed to obtain data at identical frequency points to those obtained from the higher order transform of the autocorrelated interferogram, before the division of the two functions can be performed, or a Fourier transform of identical order to that performed for the autocorrelated interferogram must be done on the zero-path difference region before the division of the two functions is obtained.

(b) For spectra of the form shown in Figures 35-2 and 35-4, the Fourier transform of the zero-path difference region of the output interferogram will be zero or negative for some portion of the computer output. The ratio of the transforms of the autocorrelated interferogram and its zero-path difference region, using interferograms as shown in Figures 35-2 and 35-4 will thus produce rubbish for the final spectra.

In general, the use of the autocorrelation technique for many and varied types of interferograms, necessitates a very complicated program to avoid the kind of troubles described above. But for all types of interferograms produced from three different multiplexing instruments, sensible output spectra have been obtained using the same program with no special case considerations being required.

Contents

36. Correlation Interferometry

R. Dick and G. Levy
Barringer Research Ltd.
Rexdale, Ontario, Canada

Abstract

A correlation interferometer for use in detecting atmospheric gases by measuring the amplitude of their interferograms is described. Results for CO, obtained using a prototype instrument, are presented.

36–1 INTRODUCTION

In recent years, much effort and ingenuity have been devoted to the improvement of methods for obtaining spectra from interferograms. As has also been pointed out by Dowling (Chapter 4) the same information is contained in the original interferogram, from which it may be extracted directly without the necessity for transformation. A case in point is the remote detection of atmospheric gases, which is of considerable importance in the field of air pollution and mineralogical exploration.

36–2 CORRELATION OF INTERFEROGRAMS

Many gases and liquids have their fundamental and most intense absorption spectra in the near- and mid-infrared, where the luminous output of thermal sources is decreasing and the available detectors are generally less efficient and noisier than those used in the visible and uv. It is therefore necessary to use the radiation as efficiently as possible, which requires instruments of large throughput. It has long been known that interferometers, by virtue of their circular symmetry, could provide an increase in throughput of the order of 100 compared with an ordinary spectrometer. Fellgett's advantage, which arises in a detector-noise-limited interferometer as a result of looking at all spectral elements simultaneously rather than sequentially, is available in the near infrared, where uncooled detectors do not attain BLIP performance. Such an interferometer can also be a very small package. In one respect, remote gas detection is simpler than spectroscopy, since it is generally known a priori which absorbers are present and one merely wishes to measure the amounts. That is, the form of the interferogram is known and only the fractional modulation is of interest. Thus, it is possible to make a system in which one compares the incoming interferogram with a reference that was made using a known amount of the target.

The method by which the interferometer is made specific to certain gases is shown in Figure 36-1, which depicts a modified Michelson Interferometer. Incoming radiation is optically filtered at F and then impinges on the beamsplitter BS, where it is amplitude divided into equal parts, which are reflected, respectively, at mirrors M_1 and M_2 before being recombined and focused on the detector D. The compensating plate C, which is of the same material and thickness as the beamsplitter, assures that both optical paths are equivalent. If mirror M_2 is displaced backwards a distance of one-quarter optical wavelength from the location at which there is zero optical path difference between the two beams, the two components of radiation, which were constructively

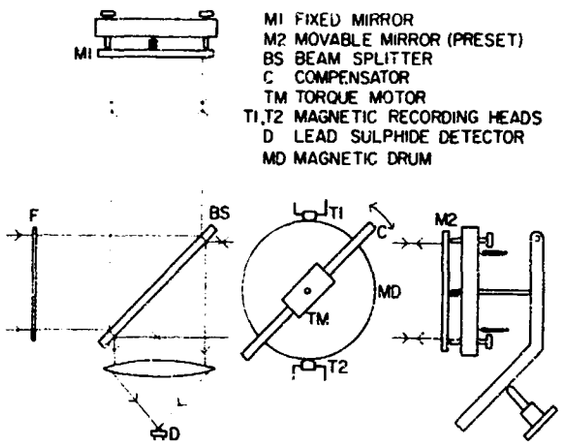


Figure 36-1. Correlation Interferometer

interfering to produce maximum radiation at D, will now destructively interfere to produce, under ideal conditions, no radiation. If the source radiation is monochromatic and M_2 moves with constant velocity, it is a simple matter to show that the detector output will be a cosine of constant electrical frequency given by $(2\nu_0 v)$ where ν_0 is the optical frequency of the input radiation in cm^{-1} and v is the velocity of displacement of the mirror in cm sec^{-1}.

Thus, it is possible to think of the interferometer as an optical analog to a superheterodyne receiver, in which each optical frequency is translated down to a much lower electrical frequency. In fact, it can be shown that the detector output, as a function of optical path difference δ, is given by

$$D(\delta)=\int_0^{\nu_M} B(\nu)(1+\cos 2\pi\nu\delta)\,d\nu, \qquad (36\text{-}1)$$

where the input radiation has power spectral density $B(\nu)$ and is band limited between 0 and ν_M.

Thus, apart from a constant term, the detector output as a function of δ is the Fourier cosine transform of the input power spectral distribution. In normal spectroscopy, where it is desired to evaluate $B(\nu)$, it is necessary to take the inverse transform of the interferogram $D(\delta)$ in order to calculate this. Such a calculation is usually performed on a computer, especially where high resolution spectra are required, and it is this computation which is the chief disadvantage of the technique. However, just as the absorption spectrum of a given gas is unique to that gas and can be used to quantitatively measure its concentration, so the Fourier transform of the spectrum is unique to the gas, and it too can be used for a quantitative measurement. That this is so is seen in Figures 36-3 to 6 which show the computed transmission spectra, modified by Gaussian limiting optical filters, for samples of methane and ammonia gas in the region of 2.3 μ and their computed transforms as a function of δ, the optical path difference.

As has been mentioned, the detector output consists basically of an electrical cosine signal at frequency $2\nu_0 v$ where ν_0 is now the center optical frequency of the entrance filter; but since the input radiation is no longer monochromatic, this cosine will be phase and amplitude modulated by the spectrum. The upper portions of both Figures 36-4 and 36-6 show the modulation envelope of this cosine as a function of δ (with scale factors of 1, 10, 100), and the lower portions illustrate the electrical angle of the cosine relative to the constantly increasing phase angle of $(\cos 2\pi\nu_0\delta)$. It is seen that both quantities are unique and quite different for the two spectra. Furthermore, since the amplitude of the transform is directly proportional to the spectral transmission of the gas, correlations made on the transform will, through suitable calibrations, yield concentrations.

Correlation is achieved as follows (Figures 36-1 and 36-2): Mirror M_2 is actually fixed and the optical path difference between the two beams is varied by using a torque motor to oscillate the compensator plate C through a small displacement about its center position. A magnetic disk is rigidly attached to the top of the oscillating refractor plate with a magnetic pick-off sensor located close to the magnetic oxide surface of the disk. Programming the instrument is described in the following paragraphs.

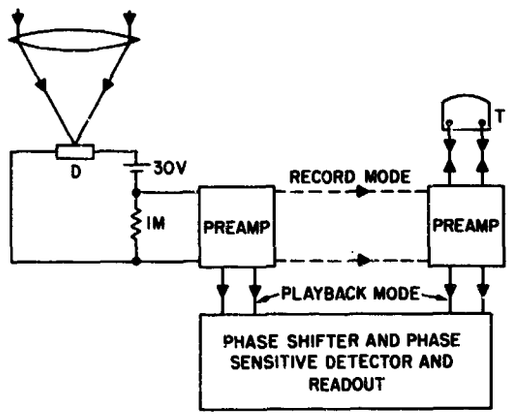


Figure 36-2. Recording and Correlation Electronics

A sample of the gas to be detected, for example CO, is inserted into the interferometer's field of view. Then, as the refractor plate is oscillated, the interferograms of the CO spectrum are generated within the interferometer to appear as a modulated sinu-

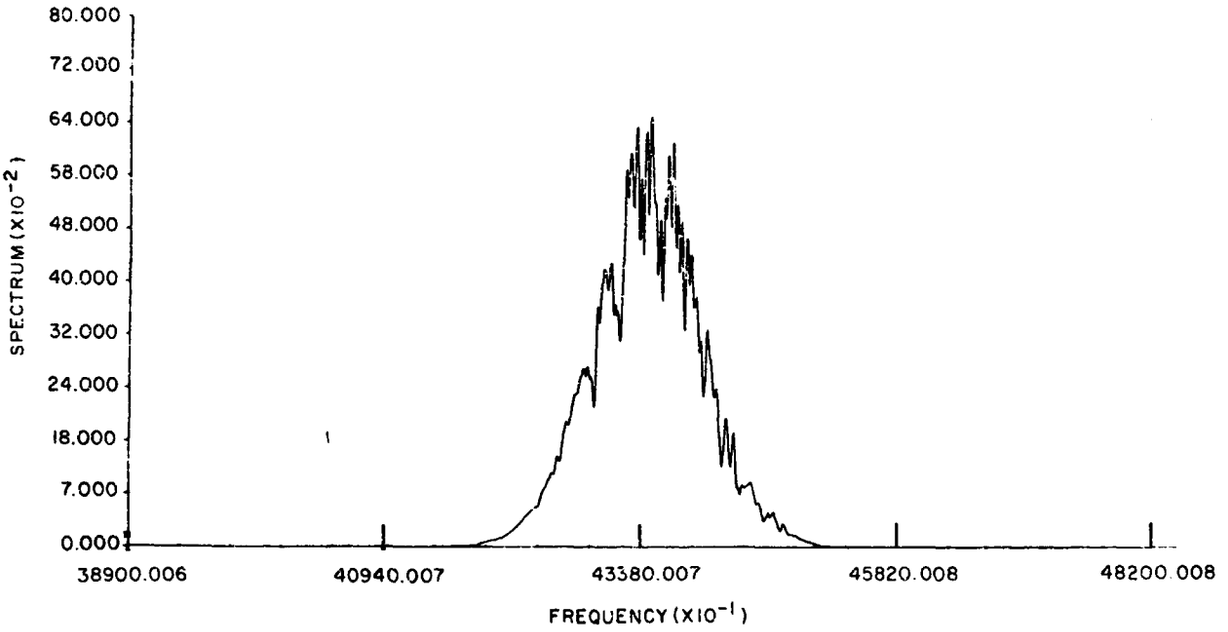


Figure 36-3. Methane Spectrum (Gaussian Filter 70 cm^{-1})

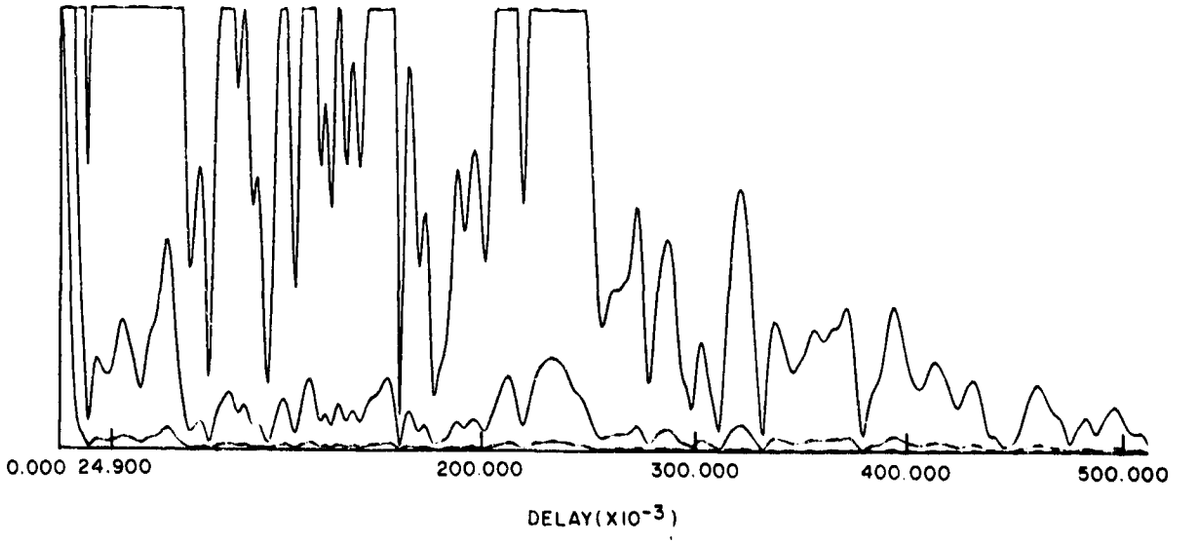


Figure 36-4. Computed Fourier Transform of Methane (0.67 atm-cm)

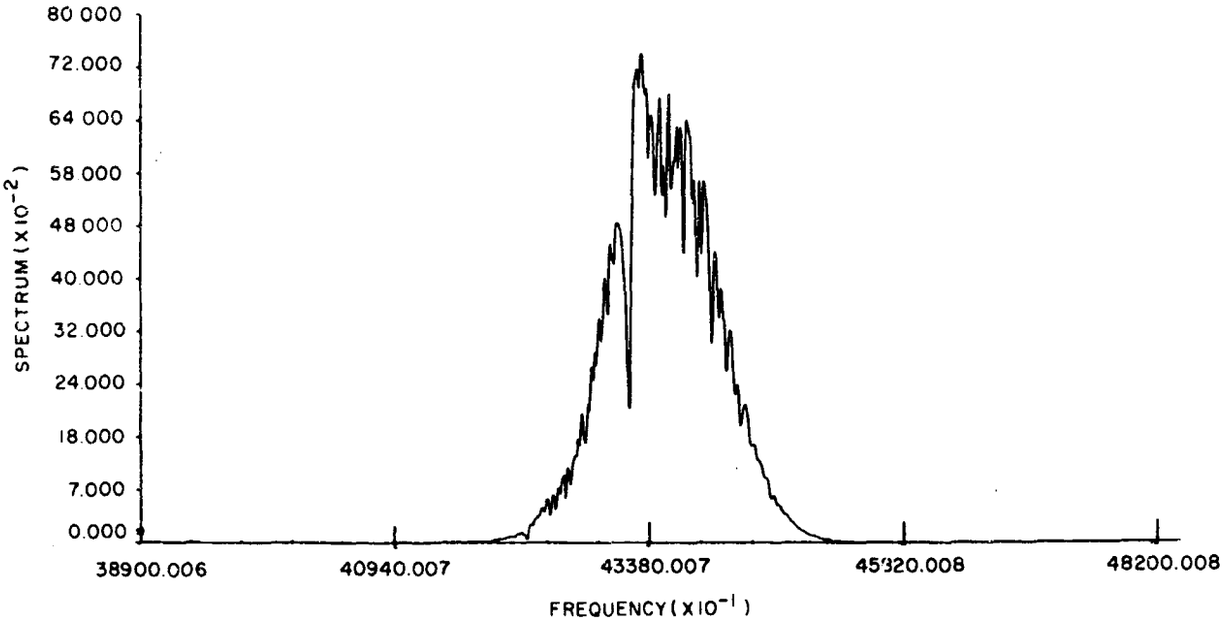


Figure 36-5. Ammonia Spectrum (Gaussian Filter 70 cm^{-1})

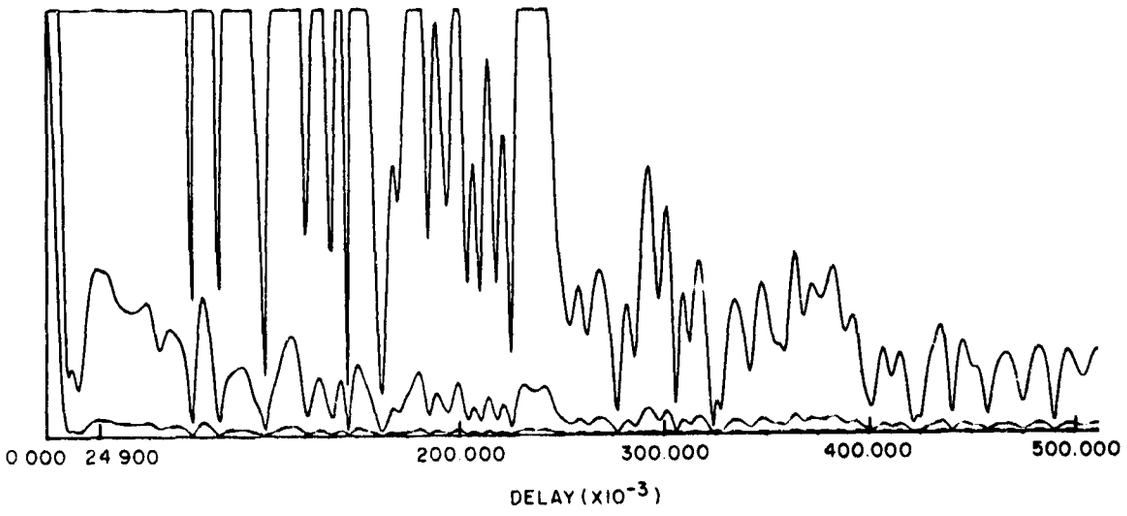


Figure 36-6. Computed Fourier Transform of Ammonia (0.73 atm-cm)

soidal output (Figure 36-7) from the IR detector. If this AC output is then coupled through to the magnetic head, the AC interferogram is recorded onto the magnetic disk.

In the remote sensing mode, the record loop is opened and the IR detector output is passed to a phase sensitive detector, which receives as its reference signal the AC playback output from the magnetic disk by way of the magnetic head. The output of the phase sensitive detector is thus a DC signal, the amplitude of which is related to the amount of gas in the instrument's field-of-view.

Assuming that an adequate path difference is taken, the transform of interfering gases will not correlate, and the output voltage will be zero in their presence. It is a simple matter to program the device for several gases by using several magnetic heads (T_1, T_2, and so forth) thus obtaining a multi-gas analyzer. Calibration of the instrument is direct, by means of the insertion of calibration cells of known concentration.

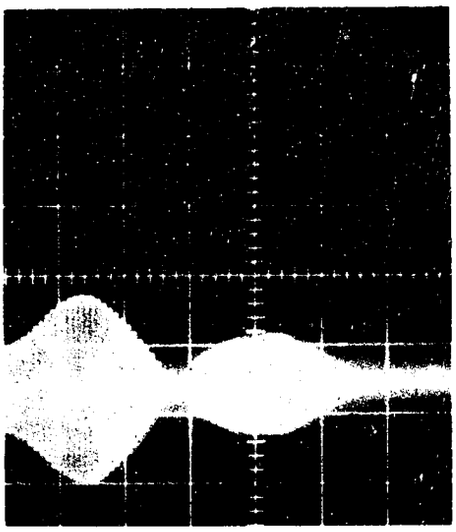


Figure 36-7. Detector Output NH_3 Near 0.5 mm PD

If the period of the sinusoid drive applied to the refractor plate by the torque motor should drift, correlation is still held, because the reference signal is a playback off the magnetic disk, rigidly fixed to the refractor plate through which the input optical signal is passing. That is, there is a fixed relation between the signal being read from the disk and the optical path difference at that instant. This gives the method an advantage over other techniques that have been used for modulation of interferometers, such as a sawtooth movement of one of the mirrors combined with fixed, narrow-band post detection filters and amplifiers for noise rejection. For example, if the mirror velocity should shift, by less than 1 percent, the effects are catastrophic. Thus, lack of uniformity or repeatability in the drive-mirror system throws out the complete wavenumber calibration of the interferometer and, more important, places narrow band detection circuits off-tune. In order to ensure reproducibility of the scan, optical feedback systems such as spectral line interferometric reference channels or moiré fringe devices are necessary. In the Barringer interferometer drift is virtually eliminated by revolving the compensator plate. Small changes in pathlength are mechanically magnified to relatively large rotations without problems of backlash, since mechanical linkages are not involved.

36-3 EXPERIMENTAL

A prototype correlation interferometer, shown in Figure 36-8, has the characteristics listed in Table 36-1. Figures 36-4 and 36-6 show that, for a specific gas, certain portions of the interferogram have a larger signal output than others, and may also have more information in the phase and amplitude variations. Thus, since only a portion of the interferogram is actually scanned, mirror M_2 is adjusted to optimize to the particular gas of interest. The instrument discussed here scanned only about 150 fringes but this would be greatly increased in an engineering model so as to achieve greater specificity. Saturation NRZ recording was used for the reference, so that the correlation was against the phase of the reference rather than the amplitude. Contact recording and readout techniques were used after preliminary tests had shown that over one million scans of the magnetic drum resulted in signal degradation of only 20 percent. The actual scan speed in operation was one scan per second. A floating head and flexible magnetic disk system has also been designed which reduces wear of the oxide coating to virtually zero.

This breadboard instrument has demonstrated detectivities for several atmospheric gases, which agree with predictions based on the absorption band strengths and the optical and detector parameters.

Figure 36-8. Breadboard Interferometer

Table 36-1. Breadboard System Parameters; Correlation Interferometer

Collecting Aperture	0.88 in. diam
Focal Length	1.9 in.
Field of View	0.045 mrad angular diam
Filter	Gaussian, 0.1 micron FWHM, centered at 2.35μ. Peak transmission 60%
Beam Splitter	Aluminum film on 6 mm fused quartz substrate. Reflection 40%, transmission 40%, absorption 20%.
Compensating Plate	6 mm fused quartz plate.
Detector	Ambient temperature lead sulphide; 0.16×0.16 in: $D^* = 3\times10^{10}$ cm $cps^{\frac{1}{2}}/W$
Correlation System	Multiplication by reference interferogram stored on magnetic tape, followed by integration ($\tau = 5$ sec)
Scan Range	150 wavelengths
Scan Rate	Scan and return in 1 sec

Extrapolating from this experience, it is quite feasible to design a system that will yield sensitivities of the order of a few ppm-m when used in a remote sensing mode with the sun as a source of infrared radiation. It is also possible, of course, to use the device as an ambient or sampling sensor, where, if a cooled detector is used, the calculated sensitivities increase to the order of one ppb over a 10-cm path length for many gases of interest.

36–4 COMPUTED INTERFEROGRAMS

As has been mentioned, each gas will produce distinctive interferogram signatures at certain characteristic path differences. In order to locate these optimum scan ranges and to predict signal-to-noise ratios for the various gases, a program has been written which uses FFT to produce interferograms from spectra given either in the form of digitized experimental data or as molecular constants. Provision is also included for limiting the spectrum by Gaussian, power series, or multilayer filter functions. Interferograms produced by this program are shown in Figures 36-3 to 6 and 36-9. Figure 36-9 shows the effect of varying temperature on the interferogram of the 4260 cm^{-1} band of CO, limited by a 70 cm^{-1} Gaussian filter. The dotted curve corresponds to a temperature of 300°K and a scale factor of × 100 and the three solid curves to 200°K with scale factors of × 1, × 10 and × 100. The strongest feature extends from 0.2 to 0.35 cm delay and corresponds to the range of spectral line separations found over the band. The relative phase varies across this region; the slope of the curve at a given delay depends on the wavelength at which the corresponding line separation occurs. It is zero for the point corresponding to the line separation at the band origin, since the phase reference was chosen as the transform of a hypothetical line at the band origin wavelength.

The basic pattern observed in the region of 0.27 cm delay is repeated at multiples of this delay; the strength of these "overtones" depends on the line shapes. As the higher overtones begin increasingly to overlap, beat phenomena are observed, first in the wings of the main pattern, but ultimately dominating the interferogram in the regions of longer delay.

The effects of temperature are seen principally in the amplitude of the interferogram. This is as expected, since the phase at short delays depends on the line positions which are almost independent of temperature, while the amplitude depends on the line intensities and thus on the temperature via the changing population of the energy levels.

36–5 CONCLUSIONS

It is possible to recover an interferogram of known form but unknown amplitude from noise, by multiplying the noisy signal by a noise-free copy of the interferogram of the correct phase. The required phase locking of the reference interferogram can be achieved by having it magnetically recorded on a medium rigidly attached to the element which produces the path difference in the interferometer.

A correlation interferometer employing these principles has been built and used to detect atmospheric gases by measuring the amplitude of their absorption interferograms.

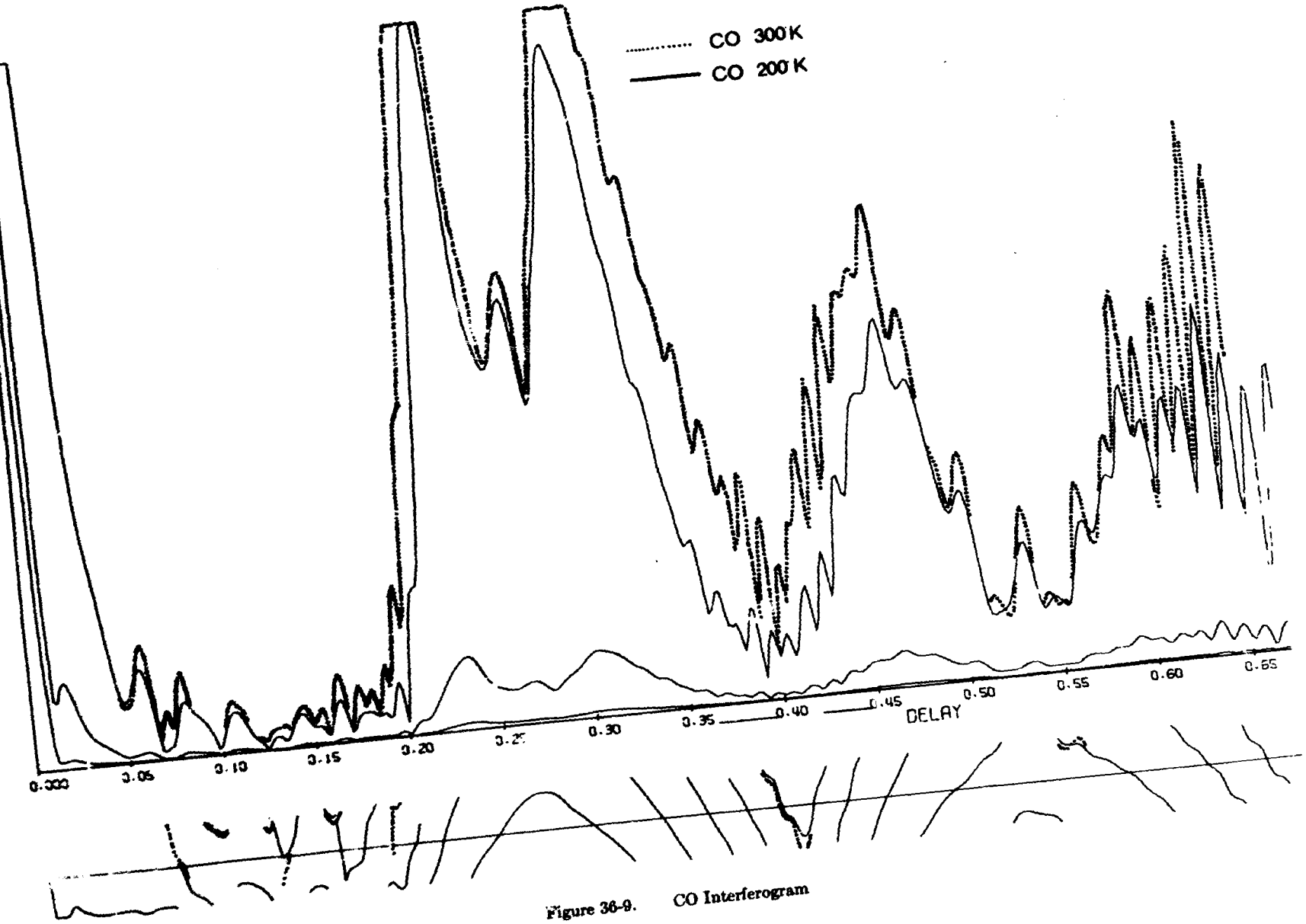


Figure 36-9. CO Interferogram

Discussion

Q. (L. W. Chaney, University of Michigan): What is the equivalent resolution of your instrument?

A. (Robert Dick): We attempt to match the resolution to the separation of the principal features of the target gas. For example, for carbon monoxide, $5\ cm^{-1}$ resolution corresponds to the average line separation.

Q. (L. N. Mertz, Smithsonian Astrophysical Observatory): Could you clarify what you mean by phase in this context?

A. (Robert Dick): We use as phase reference the cosine wave; this is the interferogram which would be produced by a sharp line at the central wavelength of the spectral interval being considered.

Q. (R. A. Hanel, Goddard Space Flight Center): Have you made an attempt to detect CO in the presence of atmospheric amounts of water vapor?

A. (Robert Dick): We have measured CO from traffic in atmospheric paths of a few hundred feet, without any evidence of H_2O interference.

Contents

37. Optical Constants of Low-Index Materials From Far-Infrared Channeled Spectra

Donald R. Smith
Air Force Cambridge Research Laboratories
Bedford, Massachusetts

Abstract

A vacuum far-infrared double-beam Michelson interferometer was used to measure the channeled spectra of several polymer films. From the channeled spectrum in the power reflectance the optical constants have been determined. A method for determining the absorption coefficient from the channeled spectrum amplitude is given and the limitations are discussed. The results of this method are compared with the results calculated from the low-resolution transmittance. The optical constants obtained for Mylar and Surlyn films in the region between 50 and 350 cm^{-1} are presented.

37–1 INTRODUCTION

The use of channeled spectra is a common method for determining the refractive index.[1,2,3] A channeled spectrum is the result of interference fringes arising from multiple internal reflections in a material. In general, a channeled spectrum is the sum of cosine terms of many orders. In Fourier spectroscopy, however, these terms can be easily separated by transforming only the appropriate part of the interferogram. If a single cosine term is isolated in this way, the channeled spectrum oscillates about zero and the index can be determined from the zero-crossings of the spectrum, since the argument of the cosine term is a function of the refractive index alone. The absorption coefficient is then determined from the low-resolution power transmittance data.

The absorption coefficient, moreover, can also be determined from the channeled spectrum, since the amplitude modulation of the channeled spectrum is a function of the absorption coefficient. Such a technique for determining the absorption coefficient will be developed. The application of this technique to reflection measurements on low-index materials is the subject of this paper. Although transmission measurements are easier to perform and can be done at normal incidence, reflection measurements are preferred for low-index materials, because of the increased visibility of the fringes. If the absorption is not small, however, the low-resolution reflectance is nearly independent of the absorption and another method of determining the absorption coefficient from the reflectance is desirable.

37–2 THEORY

37–2.1 Channeled Spectrum Method

Precise values of refractive indices can be determined using the channeled spectrum technique.

providing the absorption is not too large. The amplitude transmittance and reflectance of a plane-parallel-faced sample of thickness d and complex refractive index $\hat{n} = n + ik$, are given by the following expressions:

$$A_T = t^2\hat{a} + r^2t^2\hat{a}^3 + r^4t^2\hat{a}^5 + \cdots, \quad (37\text{-}1)$$

$$A_R = r - rt^2\hat{a}^2 - r^3t^2\hat{a}^4 - \cdots. \quad (37\text{-}2)$$

For the materials involved here $n \gg k$, and the complex nature of r and t may be ignored. For near normal incidence the polarization effects may also be neglected; then $r = (n-1)/(n+1)$ and $t = (4n)^{\frac{1}{2}}/(n+1)$ are the single surface amplitude reflectance and transmittance, respectively, and

$$\hat{a} = \exp[-2\pi\sigma\hat{n}d\cos\theta']$$

is the amplitude attentuation coefficient. Eqs. (37-1) and (37-2) may then be rewritten as

$$A_T(\sigma) = t^2\mu e^{i\delta} + r^2t^2\mu^3e^{3i\delta} + r^4t^2\mu^5e^{5i\delta} + \cdots, \quad (37\text{-}3)$$

$$A_R(\sigma) = r + rt^2\mu^2e^{2i(\delta+\varepsilon)} + r^3t^2\mu^4e^{4i(\delta+\varepsilon)} + \cdots. \quad (37\text{-}4)$$

In Eqs. (37-3) and (37-4) $\mu^2 = \exp[-4\pi\sigma kd\cos\theta']$ is the attentuation coefficient, $\delta = 2\pi n\sigma d\cos\theta'$ is the optical path difference, and θ' is the angle between the normal to the surface and the ray inside the sample.

The power reflectance $R(\sigma) = A_R(\sigma)A^*{}_R(\sigma)$ is then

$$R(\sigma) = \left[r^2 + \frac{r^2t^4\mu^4}{1-r^4\mu^4}\right] - 2r^2t^2\mu^2\left[1 - \frac{r^2t^2\mu^4}{1-r^4\mu^4}\right]\cos 2\delta + \cdots, \quad (37\text{-}5)$$

and the corresponding power transmittance $T(\sigma) = A_T(\sigma)A^*_T(\sigma)$ is

$$T(\sigma) = \frac{t^4\mu^2}{1-r^4\mu^4} + 2\left[\frac{r^2t^4\mu^4}{1-r^4\mu^4}\right]\cos 2\delta + \cdots. \quad (37\text{-}6)$$

In terms of the observed quantities the power reflectance is

$$R(\sigma) = \frac{B_s(\sigma)}{B_B(\sigma)} = \frac{\mathrm{FT}[F_s(x)]}{\mathrm{FT}[F_B(x)]}, \quad (37\text{-}7)$$

where $B_s(\sigma)$ and $B_B(\sigma)$ are the sample and reference spectra, respectively, and $F_s(x)$ and $F_B(x)$ are the sample and reference interferograms, respectively. These interferograms are treated as three separate interferograms of lengths x_0, x_1, and x_2, shown in Figure 37-1 as F_0, F_1, and F_2, respectively. Then the low-resolution reflectance is experimentally determined by taking the ratio between a sample and reference spectrum obtained by transforming only that portion of the interferogram around zero path difference (up to x_0), and is

$$R_0(\sigma) = \frac{\mathrm{FT}[F_{0,s}(x)]}{\mathrm{FT}[F_{0,B}(x)]}. \quad (37\text{-}8)$$

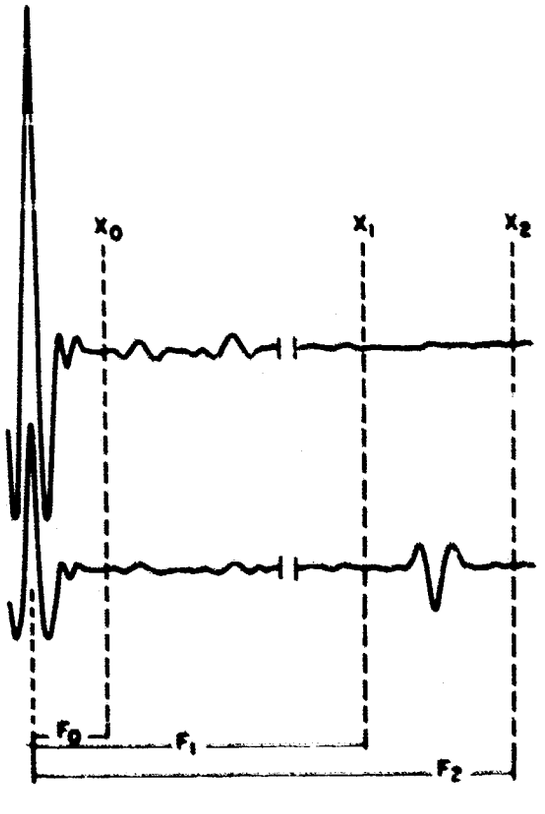


Figure 37-1. Background and Reflection Interferograms For a Material of Negligible Dispersion Showing the Points Used to Transform the Interferograms to Determine the Low-Resolution and Difference Spectra

By transforming only the portion of the sample interferogram containing the first channeled spectrum signature, that is between x_1 and x_2, and taking the ratio of this to the low-resolution reference spectrum, a single cosine term can be isolated from the channeled spectrum. This difference spectrum is related to the measured interferogram by the relation

$$R_{12}(\sigma)=\frac{\mathrm{FT}[F_{2,s}(x)]-\mathrm{FT}[F_{1,s}(x)]}{\mathrm{FT}[F_{0,B}(x)]}$$

$$=\frac{\mathrm{FT}[\Delta F_{21,s}(x)]}{\mathrm{FT}[F_{0,B}(x)]}. \qquad (37\text{-}9)$$

It should be noted here that most far-infrared materials have large enough absorption so that the channeled spectrum contains only a single cosine term to begin with, since the corresponding interferograms contain only a single signature. A result similar to Eq. (37-9) is obtained by taking the ratio of the numerator in Eq. (37-9) to the low-resolution sample spectrum. In this way the reference interferogram is eliminated entirely and the difference spectrum is

$$R_{12}'(\sigma)=\frac{\mathrm{FT}[\Delta F_{21,s}(x)]}{\mathrm{FT}[F_{0,s}(x)]}. \qquad (37\text{-}10)$$

The difference spectra in Eqs. (37-9) and (37-10) oscillate about zero and the refractive index can be determined from the zero-crossings of either spectrum. Zero crossings occur when the value of the argument of the cosine term in Eq. (37-5) satisfies the equation

$$4\pi\sigma_m nd\cos\theta'=\frac{(2m+1)\pi}{2}. \qquad (37\text{-}11)$$

The integer m is the order number corresponding to the wavenumber of the zero crossing, σ_m. Then, from Eq. (37-11), the refractive index is

$$n(\sigma_m)=\frac{(2m+1)}{8\sigma_m d\cos\theta'}. \qquad (37\text{-}12)$$

The channeled spectrum signatures must, of course, be well isolated from other modulation in the interferogram. In addition, all filters employed should be wedged to eliminate channeled spectrum signatures that may overlap the sample signature region. If a channeled spectrum signature from an unwedged filter appears in the interferogram at a distance x from zero path difference, a signature also occurs at the same distance x from the sample signature. If a filter signature is included between 0 and x_0, then the corresponding signature near the sample signature must be included between x_1 and x_2. To ensure this, choose $x_0=(x_2-x_1)/2$, where x_2 and x_1 are symmetric about the sample signature.

The observed low-resolution reflectance or unresolved channeled spectrum $R_0(\sigma)$ is simply given by the first term of Eq. (37-5)

$$R_0(\sigma)=r^2+\frac{r^2t^4\mu^4}{1-r^4\mu^4}. \qquad (37\text{-}13)$$

From Eq. (37-13) and the refractive index values calculated using Eq. (37-11), the attenuation coefficient, $u=\mu^2$, is given by

$$u=\left\{\frac{R_0(\sigma)-r^2}{r^2[t^4+r^2(R_0(\sigma)-r^2)]}\right\}^{\frac{1}{2}}, \qquad (37\text{-}14)$$

and the corresponding absorption coefficient, $\alpha=4\pi\sigma k$, is given by

$$\alpha=-\frac{\ln u}{d\cos\theta'}. \qquad (37\text{-}15)$$

If the channeled spectrum feature is well isolated, as assumed, then the channeled part of the power reflectance, $R_{12}(\sigma)$, is given by the second term of Eq. (37-5)

$$R_{12}(\sigma)=-2r^2t^2\mu^2\left(1-\frac{r^2t^2\mu^4}{1-r^4\mu^4}\right)\cos 2\delta. \qquad (37\text{-}16)$$

The solution of Eq. (37-16) results in a cubic equation in u

$$r^2u^3-r^4.1u^2-u+.1=0, \qquad (37\text{-}17)$$

where $A=-R_{12}(\sigma)/2r^2t^2\cos 2\delta$. The attenuation coefficient is the real positive root lying between 0 and 1, and the corresponding absorption coefficient is determined from Eq. (37-15).

The absorption coefficient may also be determined from the measured channeled spectrum,

$$R'_{12}=R_{12}/R_0,$$

obtained from a single interferogram.

$$R'_{12}(\sigma)=\frac{-2r^2t^2\mu^2\left[1-\dfrac{r^2t^2\mu^4}{1-r^4\mu^4}\right]\cos 2\delta}{r^2+\dfrac{r^2t^4\mu^4}{1-r^4\mu^4}}. \qquad (37\text{-}18)$$

The solution of Eq. (37-18) results in a cubic equation similar to Eq. (37-17)

$$r^2u^3+(t^2-r^2)A'u^2-u+A'=0, \qquad (37\text{-}19)$$

where $A'=-R'_{12}(\sigma)/2t^2\cos 2\delta$.

A fourth alternative is to calculate the absorption coefficient from the measured low-resolution transmittance, $T_0(\sigma)$, the first term of Eq. (37-6)

$$T_0(\sigma)=\frac{t^4\mu^2}{1-r^4\mu^4}. \qquad (37\text{-}20)$$

The attenuation coefficient is the solution of the quadratic equation

$$r^4u^2+\frac{t^2}{T_0(\sigma)}u-1=0. \qquad (37\text{-}21)$$

For materials of low refractive index, reflection measurements are to be preferred to transmission measurements because of the increased visibility of the channeled spectrum fringes for the reflection case. The visibility is defined as

$$V=\frac{I_{max}-I_{min}}{I_{max}+I_{min}}, \qquad (37\text{-}22)$$

where I_{max} and I_{min} are the maximum and minimum intensity, respectively. From Eqs. (37-5) and (37-6), the visibilities for reflection and transmission cases where $r^4\ll 1$ are

$$V_R=t^2x^2 \qquad (37\text{-}23)$$

and

$$V_T=2r^2x^2. \qquad (37\text{-}24)$$

From Eqs. (37-23) and (37-24) one finds that $V_R>V_T$ for $n<3.7$, and for this reason reflection measurements are preferred for most materials in the far-infrared region. This fact is shown quite clearly in Figure 37-2, which shows both a reflection and a

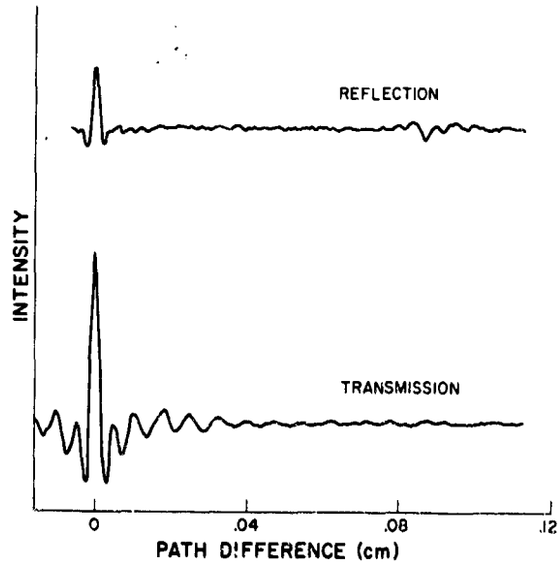


Figure 37-2. Interferograms of the Reflection and Transmission Spectra of a 251-μm Thick Mylar Sample. The channeled spectrum signature is clearly seen in reflection but not in transmission

transmission interferogram for a Mylar sample. The visibility advantage for the reflection case is obvious. Because of this advantage the refractive index of most materials can be determined with greater accuracy from reflection data than from corresponding transmission data. The situation is generally reversed where the absorption coefficient is concerned, with the most accurate results coming from the low-resolution transmittance data. For the materials presented in this paper, only the reflection interferograms showed observable channeled spectrum signatures from which the refractive index could be determined.

Since reflection measurements are required to determine n for these materials, a method for determining α from the same data is desirable. The absorption coefficient determined from the low-resolution reflectance by Eq. (37-14) is reliable for cases of small absorption only. For other cases $r^2 \gg r^2 t^4 \mu^4/(1-r^4\mu^4)$, making $R_0(\sigma) \simeq r^2$, the single surface reflectance which is independent of the absorption. Considering only the α dependent terms in Eq. (37-5), the coefficient of the cosine term is the largest quantity for all cases and should be expected to yield the most accurate values of α. This was found to be true, and the reflection results given here were calculated from the difference spectrum R_{12} using Eq. (37-18).

37-2.2 Convergence Correction For Non-Normal Incidence

Because the radiation passing through the sample is distributed over a finite solid angle, all rays do not pass through the same thickness of material in the sample. In addition, for reflection measurements the central ray of the converging cone of radiation is not incident normally to the sample but makes an angle θ_0 with the normal to the surface. Because of this, the observed relative intensity is the average over all possible angles and is

$$\langle I \rangle = \frac{\int_0^{\theta_m}\int_0^{2\pi} I_0 + I_1 \cos(4\pi n\sigma\, d \cos\gamma) + \cdots \sin\theta\, d\theta\, d\phi}{\int_0^{\theta_m}\int_0^{2\pi} \sin\theta\, d\theta\, d\phi}. \quad (37\text{-}25)$$

$$\langle I \rangle = I_0 + \frac{I_1}{\pi\theta_m^2}\int_0^{\theta_m}\int_0^{2\pi} \cos(4\pi n\sigma\, d\cos\gamma)\sin\theta\, d\theta\, d\phi, \quad (37\text{-}26)$$

where θ_m is the angle between the central ray and extreme ray of the radiation cone internal to the material, and γ is the angle of incidence of an arbitrary ray, that is, the angle between the arbitrary ray and

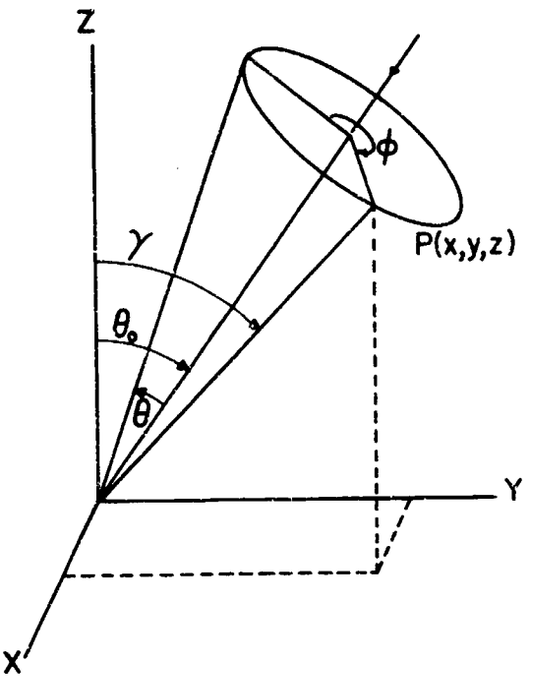


Figure 37-3. Schematic of the Conical Bundle of Radiation Converging on the Sample, Located at the Origin, Whose Surface Lies in the x-y Plane. γ is the angle of incidence of the axial ray of the bundle

the normal to the surface in the sample. Then by inspection of Figure 37-3,

$$\cos\gamma = \cos\theta_0\cos\theta + \sin\theta_0\sin\theta\cos\theta\cos\phi. \quad (37\text{-}27)$$

Applying the small angle approximation to θ, Eq. (37-27) becomes

$$\cos\gamma \approx \cos\theta_0 + \theta\sin\theta_0\cos\phi - \theta^2\frac{\cos\theta_0}{2}. \quad (37\text{-}28)$$

The above integral was treated by Lommel[4] in connection with the theory of diffraction by an aberration-free lens in 1885 and the solution is in terms of Lommel functions.* Following the notation of

* The author is indebted to G. A. Vanasse, who directed his attention to Lommel's solution to this integral just after he had completed a computer program to evaluate it numerically.

Murty[5] and Linfoot,[6] the average intensity is

$$\langle I\rangle = m\cos(4\pi n\sigma d\cos\theta_0 - \psi), \qquad (37\text{-}29)$$

where

$$m = (2/p)(U_1^2+U_2^2)^{\frac{1}{2}} \text{ for } \left|\frac{p}{q}\right|<1, \qquad (37\text{-}30)$$

$$m = \frac{2}{p}\left[1+V_0^2+V_1^2 - 2V_0\cos\left(\frac{p}{2}+\frac{q^2}{2p}\right) - 2V_1\sin\left(\frac{p}{2}+\frac{q^2}{2p}\right)\right]^{\frac{1}{2}} \text{ for } \left|\frac{p}{q}\right|>1, \qquad (37\text{-}31)$$

and

$$\psi = \left[\frac{p}{2}-\tan^{-1}(U_2/U_1)\right] \pm n\pi, \text{ for } \left|\frac{p}{q}\right|<1, \qquad (37\text{-}32)$$

$$\psi = \left[\frac{p}{2}+\tan^{-1}\left\{\frac{V_0-\cos\left(\frac{p}{2}+\frac{q^2}{2p}\right)}{V_1-\sin\left(\frac{p}{2}+\frac{q^2}{2p}\right)}\right\}\right] \pm n\pi, \text{ for } \left|\frac{p}{q}\right|>1. \qquad (37\text{-}33)$$

The Lommel functions are defined by

$$U_n = \sum_{j=0}^{\infty}(-1)^j\left(\frac{p}{q}\right)^{2j+n}J_{2j+n}(q), \qquad (37\text{-}34)$$

$$V_n = \sum_{j=0}^{\infty}(-1)^j\left(\frac{q}{p}\right)^{2j+n}J_{2j+n}(q), \qquad (37\text{-}35)$$

where J_ν is a Bessel function of order ν, and p and q are related to θ_0 and θ_m by the following relations

$$p = 4\pi n\sigma d(\cos\theta_0)\theta_m^2, \qquad (37\text{-}36)$$

$$q = 4\pi n\sigma d(\sin\theta_0)\theta_m. \qquad (37\text{-}37)$$

For normal incidence $q=0$ and the signal modulation reduces to the following expression

$$m(p,0) = \left(\frac{\sin(p/4)}{p/4}\right), \qquad (37\text{-}38)$$

which agrees with the expression for the signal modulation derived by Randall and Rawcliffe[3] for transmission measurements. The quantities m and ψ have been tabulated by Hebermehl et al.[7] for a large range of values of p and q. For the author's measurements $\theta_0 \approx \theta_m$ and therefore $p/q \approx 1$.

37-3 MEASUREMENT PROCEDURES

37-3.1 Experimental Arrangement

The channeled spectrum in reflection was measured by recording simultaneously the sample and reference interferogram with a double beam Michelson interferometer. A cooled sapphire filter was used in conjunction with black polyethylene sheets and crystal power filters to eliminate unwanted radiation. A ¼-mil Mylar beam splitter was employed to cover the spectral region between 50 and 350 cm^{-1}. The detector was a helium-cooled carbon bolometer operated at 1.5°K. The radiation was polarized by a pile-of-polyethylene-sheets polarizer. The polarization achieved with this polarizer was better than 96 percent over the entire spectral region.

37-3.2 Thickness

The thicknesses of the samples were measured by comparison with precision gauge blocks. The average thickness can be measured by this method with an uncertainty of less than one micron. The variation in thickness across the sample is typically several microns for these polymer films, and is the limiting factor in the thickness determination. The accuracy of the refractive indices is limited by the quality of the samples. The thicknesses of the Mylar and Surlyn A* samples were $251\pm1\,\mu$ and $243\pm1\,\mu$ respectively.

37-3.3 Sample Orientation

Orientation of the samples was accomplished using a polarizing microscope. The interference figures observed were those expected for biaxial crystals viewed along the acute bisectrix. With the use of a quartz wedge the optic plane was determined to be the vibration direction of the slow ray, making the "crystals" optically negative. From such samples only the slow and the medium rays can be measured, and therefore only the partial birefringence can be determined. The samples were simply cut with one edge parallel to the optic plane and were aligned in

* Surlyn A is a plastic material, called an ionomer resin, which is manufactured by E. I. DuPont de Nemours & Co.

the sample holder along the cut edges. The samples can easily be oriented to $\pm 1°$ by this method.

Dispersion of the axes and bisectrices could not be observed for any of the samples and the size of the optic angles was the same for all wavelengths in the visible, although the optic angle did vary with the thickness of the sample. (For example, for Mylar the axial angle varies in the range from 20° to 33°, depending on the thickness of the sample.) The symmetry of the color distribution caused by the dispersion in the interference figure is a reflection of the symmetry of the crystal. Since no dispersion could be observed, it is impossible to distinguish between the three possible crystallographic systems of a biaxial crystal: orthorhombic, monoclinic, and triclinic. Only in a crystal with an orthorhombic structure are the optic axes parallel to the crystallographic axes under all conditions.

37–4 MEASURED RESULTS

37–4.1 Surlyn A

A 243-μm thick sample of Surlyn A 1601 film was oriented at 45° to the optic axis to eliminate the birefringence of the sample. Using polarized light, the average refractive index and absorption coefficient were determined. The average refractive index is shown in Figure 37-4. The absorption coefficient was determined from both transmission and reflection data and is shown in Figure 37-5. The absorption coefficient values from the reflection data were calculated from the difference spectrum R_{12} by means of Eq. (37-18). The corresponding transmission values were calculated from T_0 using Eq. (37-15). Good agreement between the two methods was not achieved in every case, for reasons which are not entirely clear. The light beam in the interferometer that was used has aberrations arising from off-axis illumination of several concave mirrors. This condition can certainly produce discrepancies, but tilting of the moving mirror of the interferometer cannot be ruled out as a reason for the discrepancies.

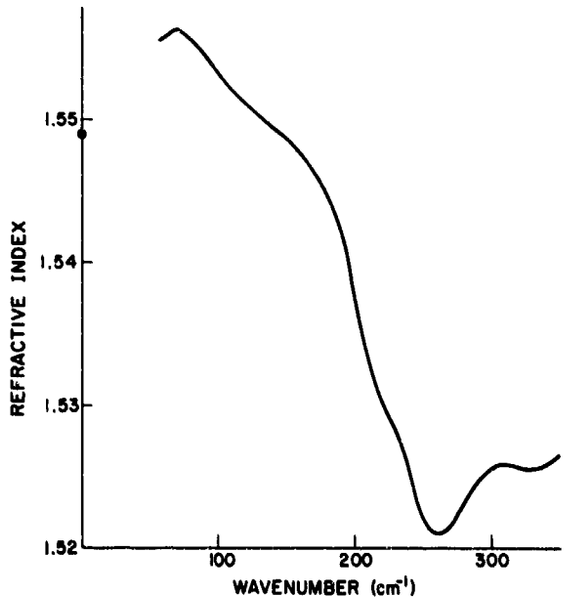


Figure 37-4. Refractive Index of Surlyn A. The circled point represents the refractive index calculated from the dielectric constant

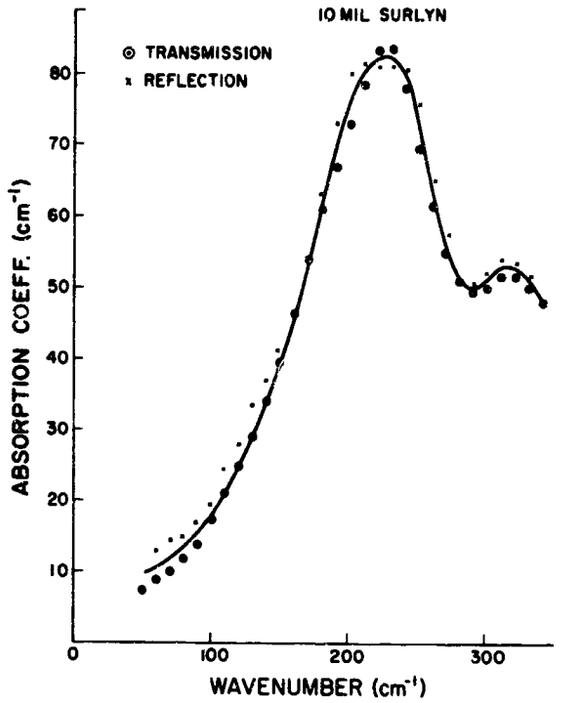


Figure 37-5. Comparison of Results for the Absorption Coefficient of Surlyn A. The circled points denote the values obtained by using the low-resolution transmittance, Eq. (37-15); the crosses denote the values obtained from reflection data using Eq. (37-18)

37–4.2 Mylar (polyethylene terephthalate)

Mylar is widely used as beamsplitter material in the far-infrared spectral region. Mylar is a partly crystalline polymer which can be treated as a mixture of crystalline and amorphous phases or as a single crystalline phase with extensive disorder and numerous defects. The structure of Mylar is similar to that of other polymers where portions of many molecules are packed side by side in a precise crystalline fashion, each molecule passing through several crystalline regions. The amorphous regions consist of portions of chain molecules which tie one crystalline region to the next.

Drawing of polymers which are semicrystalline produces orientation of crystallites, and orientation of molecules in the non-crystalline portions of the polymers in the direction of draw. Orientation perpendicular to the direction of draw is random. Rolling, on the other hand, may orient the molecules both perpendicular and parallel to the direction in which the polymer

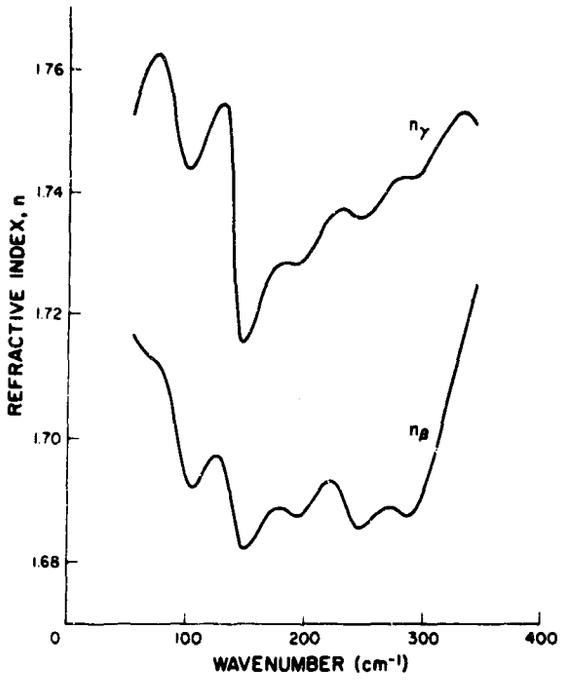


Figure 37-6. Slow and Medium Ray Refractive Indices of Mylar in the 50- to 350-cm^{-1} Spectral Region as Determined From a 10-Mil type A Film. β and γ are the medium and slow rays for visible light, respectively

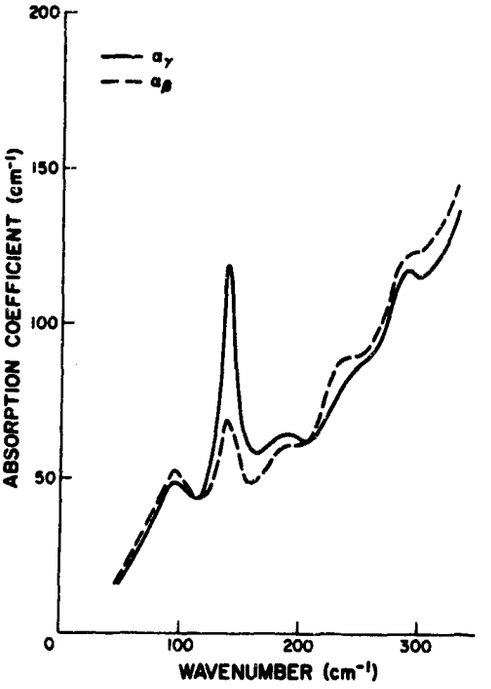


Figure 37-7. Slow and Medium Ray Absorption Coefficient of Mylar

was rolled. In such cases of double orientation the polymer is optically biaxial. Hot drawn fibers of Mylar give sharp x-ray diffraction patterns, showing that they are highly crystalline and that the crystalline regions are well oriented. The proportion by weight of crystalline regions in drawn fibers is from 48 to 65 percent or higher.[8] From these x-ray diffraction patterns the unit cell has been determined to be triclinic. Although the symmetry of the individual crystallites in a polymer may be quite low, in fibers and films containing large numbers of small crystallites a kind of pseudo-symmetry is introduced by the way in which they are distributed. Only if this pseudo-symmetry of the bulk material is orthorhombic, are the optic axes the same for all wavelengths.

A 251-μm-thick sample of type A film was oriented with the use of a polarizing microscope. From this sample the slow and medium ray refractive indices were determined and are given as n_γ and n_β, respectively, in Figure 37-6. The absorption coefficients corresponding to the slow and medium rays are shown in Figure 37-7. Since the absorption coefficients determined from the reflectance and the transmittance data differed by less than 10 percent over the entire spectral range, the results shown represent the average of the two sets of data. In an earlier study on Mylar by Bell,[9] only the spectral phase and power transmittance were presented. The optical constants from these data were calculated, and the results were found to be in reasonable agreement with the values presented, although Bell's sample orientation and polarization conditions are unknown.

37-5 SUMMARY

A method for determining the refractive index of materials from the channeled spectrum has been presented. It has been shown that reflection measurements are preferred for most materials in the far infrared. The methods of determining the absorption coefficient from a reflection channeled spectrum have been discussed and a technique for determining α from the amplitude of the fringes has been presented. The major problem in the determination of the absorption coefficient from the fringe amplitude has been found to be attenuation of the modulation in the interferogram due to misadjustment of the interferometer and associated optics. Attenuation of the channeled spectrum signature gives rise to a false absorption which is attributed to the sample. The absorption coefficient determined from the low resolution transmittance or reflectance is relatively unaffected by the effects of a misadjustment of the interferometer. These methods have been used to determine the optical constants of a number of materials in the 50- to 350-cm^{-1} spectral region and results are presented on two samples to illustrate the technique. The results for α have been compared with those obtained from the low-resolution transmittance data and have been found to be in good agreement.

References

1. Roberts, S., and Coon, D. D. (1962) *J. Opt. Soc. Am.* **52**: 1023.

2. Loewenstein, E. V. (1961) *J. Opt. Soc. Am.* **51**: 108.

3. Randall, C. M., and Rawcliffe, R. D. (1967) *Appl. Opt.* **6**: 1889.

4. Lommel, E. (1885) *Abhandl. bayer. Akad. Wiss. Math.-Naturw. Kl.* **53**: 233.

5. Murty, M. V. R. K. (1960) *J. Opt. Soc. Am.* **50**: 7.

6. Linfoot, E. H. (1955) *Recent Advances in Optics*, Oxford University Press, New York, pp. 37, 38.

7. Hebermehl, G., Minkwitz, G., and Schulz, G. (1965) *Tabellen Der Lommelschen Funktioner* $U_1(w, z)$ und $U_2(w, z)$ *Zweier Veränderlicher Sowie Abgeleiteter Funktionen* Akademie-Verlag, Berlin.

8. Daubeny, R. De P., et al. (1954) *Proc. Roy. Soc.* **A226**: 531.

9. Bell, E. E. (1966) *Infrared Phys.* **6**: 57.

Discussion

Q. (James L. Lauer, Sun Oil Co.): What is the lowest refractive index you have been able to measure or believe can be measured by your technique?

A. (D. Smith, AFCRL): I am not sure. We have measured materials with refractive index of about 1.4, but I do not know what the lower limit for this method is.

Contents

38. Interferometric Measurements of the Far-Infrared Refractive Index of Sodium Fluoride at Low Temperatures

C. M. Randall
Space Physics Laboratory
The Aerospace Corporation
El Segundo, California

Abstract

The refractive index of sodium fluoride at cryogenic temperatures has been determined between 30 and 180 cm^{-1} from channel spectra obtained by Fourier spectroscopy. By fitting the measured refractive index to a Lorentz oscillator model, a low-frequency dielectric constant of 4.70 and a damping parameter varying from 3.3 cm^{-1} at 80° K to less than 0.025 cm^{-1} at 10° K are found.

38–1 INTRODUCTION

Fourier spectroscopy permits the measurement of optical constants in the far-infrared spectral region with an accuracy rivaling that obtained by specialized instruments in other spectral regions. These techniques have been applied to the determination of the far-infrared refractive index of sodium fluoride at low temperatures. First the technique employed is briefly reviewed, the experimental problems of applying the method to low temperature measurements are then discussed, and finally the optical constants obtained for sodium fluoride at liquid nitrogen and liquid helium temperatures are presented.

38–2 CHANNEL SPECTRUM METHOD

The Fourier spectroscopic techniques employed for refractive index measurements may be divided into two classes, which depend on the location of the sample with respect to the interferometer. In the first class the sample is located within one beam of the interferometer (asymmetric Michelson) and so both phase and amplitude information are available, from which both real and imaginary parts of the refractive index may be determined. In the second class the sample is located outside the interferometer (channel spectrum) and so the real part of the refractive index can be determined with high accuracy, based on wavenumber measurements, while a less accurate determination of the imaginary part (absorption) is possible based on spectral intensity measurements. Since the two interfering beams in the Aerospace lamellar grating interferometer[1] are not conveniently separated spatially, the channel spectrum method was used. [This work was conducted under U. S. Air Force Space and Missile Systems Organization (SAMSO) Contract F04701-69-C-0066].

The basic sample geometry is shown in Figure 38-1. For light incident on a sample of thickness h at any angle, a straightforward and familiar calculation

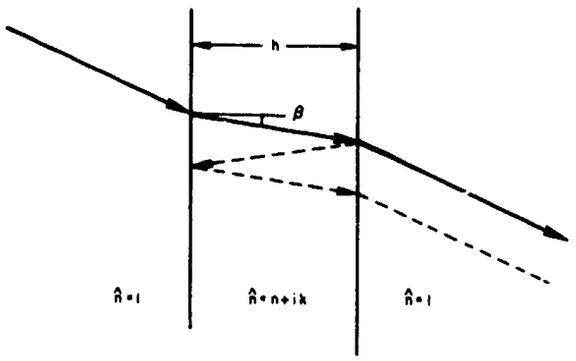


Figure 38-1. Channel Spectrum Sample Geometry

shows that the transmittance $T(\nu)$ is

$$T(\nu)=\tau(\nu)(1+2\sum_{l=1}^{\infty}\rho^l\cos l\theta), \qquad (38\text{-}1)$$

where $\nu(=1/\lambda)$ is the wavenumber and the other quantities are related to the complex refractive index $\hat{n}=n+ik$ by the following relations[2]

$$\tau(\nu)=16(n^2+k^2)\exp(-\alpha h\cos\beta)/\{[(n+1)^2+k^2]^2(1-\rho^2)\}, \qquad (38\text{-}2)$$

$$\rho=[(n-1)^2+k^2]\exp(-\alpha h\cos\beta)/[(n+1)^2+k^2], \qquad (38\text{-}3)$$

$$\theta=4\pi h n\nu\cos\beta+\tan^{-1}[2k/(n^2+k^2-1)], \qquad (38\text{-}4)$$

and $\alpha=4\pi\nu k$ is the customary absorption constant; β is the angle between the ray and the normal within the sample. The complete transmittance spectrum given by Eq. (38-1) consists of a slowly varying average transmittance given by $\tau(\nu)$ modulated by the interference given by the cosine series. The absorption parameter α, may be obtained from the average transmittance by an iterative solution of Eq. (38-2), once n is known. The modulation will have maxima when the optical thickness is an integral number of half-wavelengths; that is, when $\theta=2m\pi$, where m is called the order number. If ν is the wavenumber at which a maximum occurs, the real part of the refractive index, which can be obtained from Eq. (38-4), is

$$n=m/(2\nu h\cos\beta). \qquad (38\text{-}5)$$

For the samples studied the imaginary part of the index is small enough to allow the arc-tangent term to be completely ignored.

The assignment of order number m must be made with particular care in a material with significant dispersion, such as sodium fluoride. This assignment is made by assuming a functional form for the dependence of n on frequency, and then solving Eq. (38-5) simultaneously for two different maxima at as low a frequency as possible to obtain a value for the refractive index. This value of n is then used to compute the order number from Eq. (38-5). Since the spectral range of the data is well below the characteristic lattice absorption frequency, ν_0, it is assumed that a series expansion of Sellmier's equation is an adequate function to relate the real refractive index to ϵ_0 and ϵ_∞, the low and high frequency dielectric constants, respectively, as follows

$$n=n_0\{1+A(\nu/\nu_0)^2[1+B(\nu/\nu_0)^2]\}, \qquad (38\text{-}6)$$

where $A=(\epsilon_0-\epsilon_\infty)/2\epsilon_0$, and $B=(3\epsilon_0+\epsilon_\infty)/4\epsilon_0$. A and B may be calculated from crude estimates since they affect only small corrections; $n_0=\epsilon_0^{\frac{1}{2}}$ will be determined from the simultaneous equation solution. By substituting the value of n_0 back into Eq. (38-5) the order number, correct to first order in $(\nu/\nu_0)^2$, is found after some algebra, in terms of only measurable quantities and A, to be

$$m=(\Delta m/\Delta\nu)\{1-A(\nu/\nu_0)^2[2+3\phi+\phi^2]\}, \qquad (38\text{-}7)$$

where $\Delta\nu/\Delta m$ is the frequency change between maxima in the region of the maximum of order m at frequency ν, and $\phi=\Delta\nu/\nu$. In the worst case the correction term was about 4 percent and resulted in a change of 4 in the order number. Of course, once the order number is correctly assigned at one place in the spectrum it is assigned correctly for the entire spectrum.

The form of Eq. (38-1) as a cosine series emphasizes the advantage of taking data for this method by means of an interferometer which Fourier cosine transforms Eq. (38-1) so that the coefficients of the various terms become physically separated in "signatures" occurring at regularly spaced intervals in interferometer optical path difference. This is taken advantage of by obtaining data away from zero path difference in such a manner as to maximize the signal-to-noise ratio in the signatures. This optimizes the precision of the wavenumber determination in the modulation of the transmittance spectrum. The portion of the interferogram about zero path difference is determined in such a manner as to optimize the amplitude accuracy

of the spectrum. The average transmittance is determined from this amplitude spectrum appropriately corrected. Room temperature refractive index measurements by these methods have already been completed for a number of materials.[2,3]

38-3 LOW TEMPERATURE MEASUREMENTS

These measurements have been extended to low temperatures because this gives increased insight into the mechanisms for optical constant change in solids and because many potential uses of materials in the far-infrared will be in cooled systems.

Because of the convenience of maintaining only one cryostat, many investigators place their samples in the light pipe leading from the interferometer to the cooled detector. For the work described here it was decided to use the more complicated route of a separate sample dewar to permit more careful definition of the angles with which radiation strikes the sample, and to allow the sample temperature to be varied independently while the detector is maintained at its optimum operating temperature.

With a separate dewar it is important that differences in emission between a cold sample and its warm surroundings should not be confused with changes in transmittance. Figure 38-2, which is a

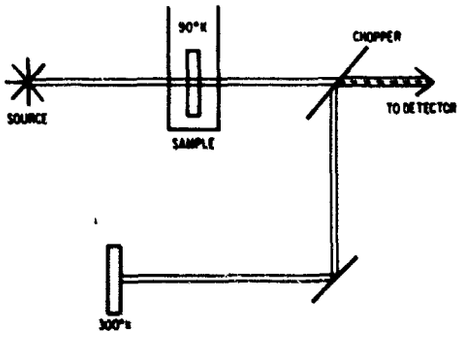


Figure 38-2. Schematic Diagram of the Aerospace Lamellar Grating Interferometer Sample Optics Arranged For a Single-Beam Run

schematic diagram of the sample optics of the Aerospace interferometer, illustrates how such confusion may arise. In an absorbing spectral region the transmittance of the sample is low and the emissivity is high, so the detector sees the difference between a 300° K object in the lower beam and a 90° K object in the upper beam. The effect is to superimpose on the transmittance spectrum a spectrum of the difference in emission. Figure 38-3 shows such a difference-of-emission spectrum, comparing polyethylene at liquid nitrogen temperature with black anodized aluminum at room temperature. No source

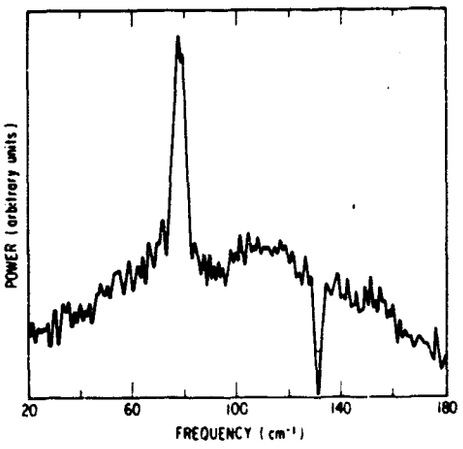


Figure 38-3. Spectrum Due to the Difference in Emission Between a Black Anodized Aluminum Sheet at Room Temperature and Polyethylene at Liquid Nitrogen Temperature. The detecting electronics are adjusted so that increasing emission from the aluminum gives a positive signal. The absorption at 130 cm^{-1} is due to a cold crystal quartz window in the detector cryostat

illuminates either material. Over most of the spectral region the detector sees through the quite transparent polyethylene to the room temperature surroundings, and so the difference spectral intensity is not large. In the polyethylene absorption band at 70 cm^{-1}, however, the detector views an object at liquid nitrogen temperature. The room temperature alumi-

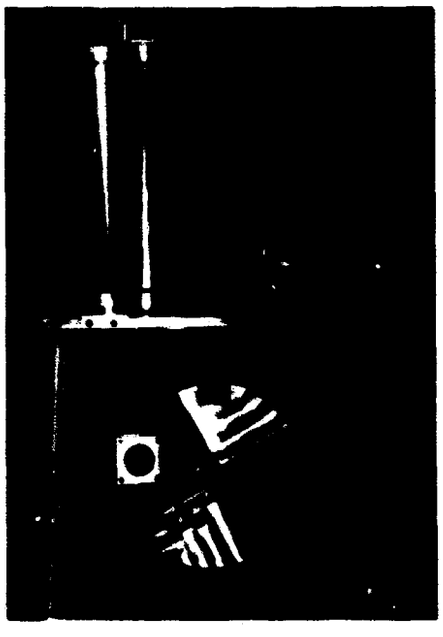


Figure 38-4. Dewar and Chopper for the Aerospace Lamellar Grating Interferometer

num emits more energy than the polyethylene, so the difference spectral intensity increases significantly. This occurs at just the places in the spectrum where there is usually the most interest in obtaining accurate measurements of transmittance from the composite spectrum.

These problems may be avoided by chopping the radiation before it passes through the cold sample. This has required the installation of a new chopper in the Aerospace interferometer. It is shown, along with the dewar, in Figure 38-4. This chopper is driven by a stepping motor which allows a variation in chopping frequency from zero up to about 25 Hz. The samples are mounted behind the window in the dewar tail on a high-conductivity copper cold finger.

38-4 SODIUM FLUORIDE RESULTS

Sodium fluoride was chosen for study for several reasons. There is considerable theoretical interest in the simple cubic ionic crystals. The alkali halides show a super-transparency at low temperatures, indicating marked change in their optical parameters. Also, there is some diversity in the values quoted for the low-frequency dielectric constant.[4] Interferograms have been obtained of the channel spectrum and the average transmittance spectrum of a 0.5639-cm-thick crystal of sodium fluoride at liquid nitrogen and liquid helium temperatures.

Because of the dispersion in NaF the signatures in the interferogram which give the channel spectrum are very asymmetric. This is emphasized in Figure 38-5, which shows portions of the NaF signature and,

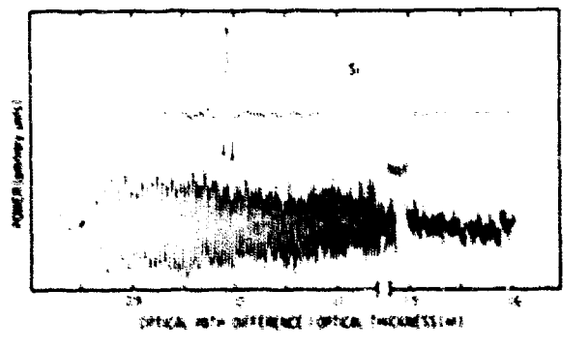


Figure 38-5. Channel Spectrum Signature of Sodium Fluoride at Liquid Helium Temperature. The signature of a silicon sample is shown for comparison

for comparison, the signature of a sample of very low-dispersion silicon.

Figure 38-6 shows the refractive index obtained by the means outlined above from the Fourier cosine transform of interferograms such as that shown in Figure 38-5. To determine the low-frequency dielectric

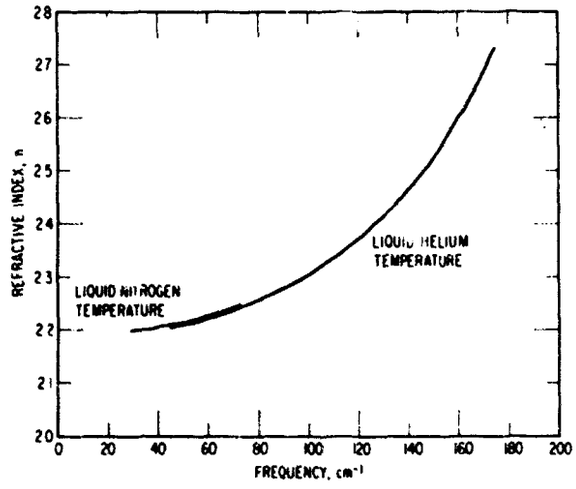


Figure 38-6. Real Part of the Refractive Index of Sodium Fluoride at Liquid Nitrogen and Liquid Helium Temperatures

constant ϵ_0 more accurately, the real refractive index data have been least-squares fit to a Lorentz oscillator model for which the complex dielectric constant, $\epsilon = (\hat{n})^2 = (n+ik)^2$, is

$$\epsilon = \epsilon_\infty + (\epsilon_0 - \epsilon_\infty)/[1-(\nu/\nu_0)^2 + i(\gamma\nu/\nu_0^2)], \quad (38\text{-}8)$$

where ϵ_∞ is the dielectric constant at frequencies far above the characteristic frequency, ν_0, and γ is a damping constant. The values chosen for the con-

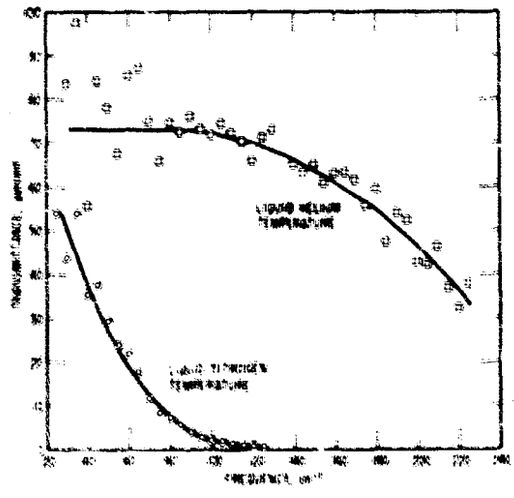


Figure 38-7. Far-infrared Transmittance of Sodium Fluoride at Low Temperatures. O liquid nitrogen temperature, □ liquid helium temperature

stants are $\epsilon_\infty = 1.755$ and $\nu_0 = 246\ \text{cm}^{-1}$.[5,6] The γ is obtained from the average transmittance spectrum analysis to be discussed. Only ϵ_0 is allowed to vary. The best fit value is 4.70 ± 0.02. No significant change is observed with temperature in the real part of the refractive index in the spectral region where data from both liquid nitrogen and liquid helium temperatures are available.

There is, however, a significant decrease in the imaginary part of the refractive index with temperature as indicated by the temperature dependence of the average transmittance shown in Figure 38-7. The curves shown in Figure 38-7 were made by transforming only the portion of a single beam interferogram near zero path difference and correcting for source spectral variation, window transmittance, and sensitivity changes of the instrument to obtain the average transmittance $\tau(\nu)$. From these transmittance curves, the absorption, $\alpha = 4\pi\nu k$, was found by means of Eq. (38-2) and the measured real refractive index. The resulting absorption parameters are shown in Figure 38-8.

To characterize this change in absorption with temperature by a single parameter, the Lorentz oscillator has been used in a least-squares fit to the measured absorptions, with only γ varying. The value chosen for ϵ_0 was 4.70 from the real index measurements, and the other constants were the same as those used in the real index calculations. For the liquid nitrogen temperature data this procedure was straightforward, and a value of $3.3\ \text{cm}^{-1}$ was obtained for γ. For comparison, the calculated absorption for a Lorentz oscillator with this damping is shown in Figure 38-8. At liquid helium temperature the absorption is so small that the noise in the limited data available obscures any clear trend in the absorption. Therefore, the absorption for several values of γ, which give absorption parameters that bracket most of the experimental points, have been plotted in Figure 38-8.

Values of the damping constant at intermediate temperatures will be of interest to differentiate between various mechanisms suggested for the absorption of infrared radiation by the alkali halides, since different mechanisms have somewhat different absorptance temperature dependence. The dewar is now being modified for variable temperature operation, so that this functional dependence on temperature may be investigated.

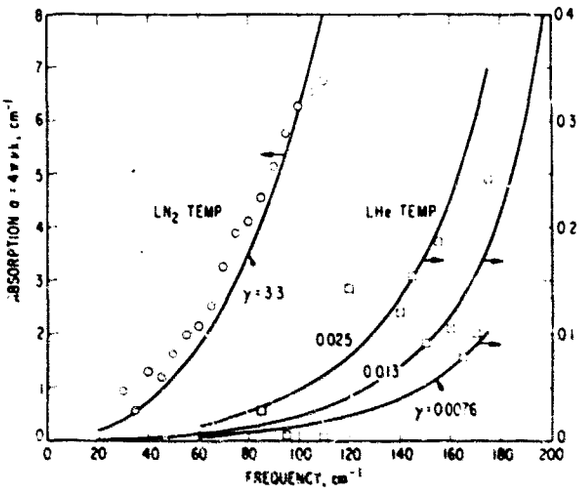


Figure 38-8. Absorption $\alpha = 4\pi\nu k$, for Sodium Fluoride at Low Temperature. ○ liquid nitrogen temperature, □ liquid helium temperature. The smooth curves are absorptions calculated for a Lorentz oscillator model dielectric with $\epsilon_0 = 4.70$, $\epsilon_\infty = 1.755$ and $\nu_0 = 246.0\ \text{cm}^{-1}$ and the values of γ in units of cm^{-1} shown on the figure

Acknowledgments

It is a pleasure to acknowledge the contributions of Drs. J. M. Dowling, R. T. Hall, and R. D. Rawcliffe to this work through frequent discussions, and the contribution of T. E. Mott in carrying out the experimental work.

References

1. Hall, R. T., Vrabec, D., and Dowling, J. M. (1966) *Appl. Opt.* **5**: 1147.

2. Randall, C. M., and Rawcliffe, R. D. (1967) *Appl. Opt.* **6**: 1889.

3. Randall, C. M., and Rawcliffe, R. D. (1968) *Appl. Opt.* **7**: 213.

4. Buyers, W. J. L. (1967) *Phys. Rev.* **153**: 923.

5. *American Institute of Physics Handbook*, Second Edition (1963) D. E. Gray, Ed., McGraw Hill, New York.

6. Randall, C. M., Fuller, R. M., and Montgomery, D. J. (1964) *Solid State Commun.* **2**: 273.

Discussion

Q. (L. Leopold, Georgetown University): In your expression for the phase there appeared a term δ, which expresses the phase shift due to absorption. You mentioned that this is negligible compared to the first term. How do you treat δ when the first term is equal to $2n^2$?

A. (C. Randall, Aerospace Corp.): Since frequencies corresponding to transmittance maxima ($\theta = m2\pi$) are measured, and since δ is only indirectly and weakly dependent on frequency (see Eq. (38-4) where δ is explicitly defined), if this term were important it would have the effect of making the order number m appearing in Eq. (38-5) slightly different than an integer.

Q. (L. Thorpe, Beckman Research & Industrial Instruments): Could you indicate the problems, and their solutions, involved in the measurement of the thickness of the parallel-sided samples used in your determinations of their refractive index as a function of frequency and temperature?

A. (C. Randall): Assuming the order number has been chosen correctly, the accuracy of the real refractive index determination depends only on the accuracy of the wave number determination and the thickness measurement. Because of uncertainties in the thermal expansion coefficients for NaF we have determined the thickness with a micrometer accurate to only ± 0.0001 in. More accurate thickness measurements have been required for room temperature work. See Randall and Rawcliffe[2] for details of these measurements as well as a discussion of a correction to sample thickness required by the convergence of the beam.

Contents

39. A New Field in Fourier Spectroscopy: Interferometric Polarimetry*

A. L. Fymat and K. D. Abhyankar†
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Abstract

The possibility of measuring the intensity and state of polarization of optical radiation by means of the high-resolution Fourier spectroscopic method is demonstrated. In the proposed experimental arrangement, a two-beam interferometer is used with a polarizer in each beam. After recombination the emergent radiation is analyzed with a linear polarizer. It is shown that the interferograms obtained in this way contain information about the four Stokes parameters of the incident radiation. The polarizers introduce an asymmetry in the interferograms requiring full (exponential) transforms for retrieval of the desired data. The effects of the finite range of path difference and the variation of its zero point with frequency are considered, and evaluation of the corresponding phase error with a proper choice of the polarizer settings is discussed. The formalism also takes into account the residual polarization introduced by the beam splitter, and the differential transmission of the two beams. Generally, three independent interferograms are needed for determining the phase error and the four Stokes parameters. Some simple arrangements are described in which the two beams are both linearly polarized.

39–1 INTRODUCTION

During the past two decades, interferometric (Fourier) spectroscopy, particularly in the infrared region of the spectrum, has been made possible by the works of Fellgett,[1] Jacquinot,[2] Connes,[3] and Lowenstein.[4] The well-known advantages of Fourier spectroscopy have so far been used only for measuring the brightness spectrum. It is the purpose of the present paper to demonstrate that the Fourier spectroscopic technique can be adapted for the measurement of all polarization parameters, that is, brightness, degree of polarization, orientation of the plane of polarization and ellipticity, with a comparably high degree of resolution over as wide a spectral range. In comparison, the conventional methods of polarimetry neither possess such a resolution nor yield a sufficient signal-to-noise ratio for the same resolution. The high-resolution measurements of both the brightness and

* The results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract No. NAS 7-100, sponsored by the National Aeronautics and Space Administration are presented.

† Senior Postdoctoral Resident Research Associate of NRC-NASA. On leave of absence from Osmania University, Hyderabad, India.

state of polarization are needed in many physical problems such as the interpretation of planetary radiation. In particular, the variation of polarization across a band in the planet's spectrum will provide valuable information on the scattering properties and structure of its atmosphere. It is here that the proposed Fourier spectroscopic method will find one of its most urgent applications.

39-2 POLARIZATION PARAMETERS AND METHODS OF MEASUREMENT

The basic representation of the state of polarization of a quasimonochromatic radiation field given by the column matrix,

$$E=\begin{pmatrix}E_x\\E_y\end{pmatrix},$$

is in terms of the time averages (denoted by angular brackets) of products of the complex field components $\langle E_i E_j^*\rangle = J_{ij}$, $(i=x,y;j=x,y)$, where the asterisk represents the complex conjugate. They define the coherency matrix (compare, for example, Wolf[5]),

$$J=\langle E\times E^\dagger\rangle=\begin{pmatrix}J_{xx}&J_{xy}\\J_{yx}&J_{yy}\end{pmatrix}, \tag{39-1}$$

where $E^\dagger$ denotes the Hermitian conjugate of E and the cross represents the Kronecker product. The elements of J are related to the Stokes parameters of the field in the following way:

$$\begin{cases} I=J_{xx}+J_{yy}=\text{Trace } J\\ Q=J_{xx}-J_{yy}\\ U=J_{xy}+J_{yx}\\ V=-i(J_{xy}-J_{yx})\end{cases} \tag{39-2}$$

It may be noted that J_{xx} and J_{yy} are real quantities whereas J_{xy} and J_{yx} are complex conjugates of each other. It is convenient to write $J_{xy}=|J_{xy}|\exp[i\beta_{xy}]$. Hence, all the Stokes parameters are real and have the dimensions of intensity.

The effect of any optical device on the electric field can be represented by a 2×2 Jones matrix K such that[6]

$$E'=KE. \tag{39-3}$$

The coherency matrix of the resultant field, J', can be obtained from E' with the help of Eq. (39-1). The emergent intensity given by the trace $Tr\,J'$ of the matrix J' is a function of the Stokes parameters of the incident field. For example, if the light is passed through a compensator which retards the phase of the y-component with respect to that of the x-component by an angle ϵ, and then through a linear polarizer making an angle θ, with the positive x-direction, the emergent intensity will be given by the interference equation[7]

$$I(\theta,\epsilon)=J_{xx}\cos^2\theta+J_{yy}\sin^2\theta+|J_{xy}|\sin 2\theta\cos(\beta_{xy}-\epsilon). \tag{39-4}$$

It is obvious from Eq. (39-4) that four independent measurements of $I(\theta,\epsilon)$ for various combinations $\{\theta,\epsilon\}$ will be sufficient for determining the four Stokes parameters. Born and Wolf[7] have suggested the following six measurements: $\{0°,0°\}$, $\{45°,0°\}$, $\{90°,0°\}$, $\{135°,0°\}$, $\{45°,90°\}$, and $\{135°,90°\}$. However, there is a redundancy in these measurements because the second and the fourth setting, and the fifth and the sixth setting, are not independent.

The usual optical methods for determining polarization are based on Eq. (39-4) and involve only measurements of two to four intensities with the help of linear polarizers and fixed retarders. On the other hand, in radio astronomy, in addition to these methods, other procedures based on the coherency concept are also used.[8,9,10] In the latter, only J_{xx} and J_{yy} are measured directly as intensities using linear or circular antennas, whereas $|J_{xy}|$ and β_{xy} are obtained by means of electronic correlation techniques involving a continuously variable retardation between the two components. The variable retardation is introduced either with a phase-sweeping device connected to a single radio telescope or with a two-beam radio interferometer. Such approaches have not, however, been made in the optical region. It is clear that the phase-sweep technique will not be possible at high optical frequencies, but the interferometric method can certainly be used as explained in the following paragraphs.

39-3 PRINCIPLE OF THE EXPERIMENTAL SCHEME

Figure 39-1 is a schematic diagram of the proposed two-beam optical interferometer to be used for measuring polarization. The radiation from the source O is first divided into two beams with the help of a splitter, which could be in the form of two slits of Young's experiment, the beam splitter of a Michelson interferometer, or any other similar arrangement. Let S_1E_σ and S_2E_σ be the complex amplitude vectors at wavenumber σ for the two beams, where S_1 and S_2 are the Jones matrices representing the action of the beam splitter. The latter include both the fractional trans-

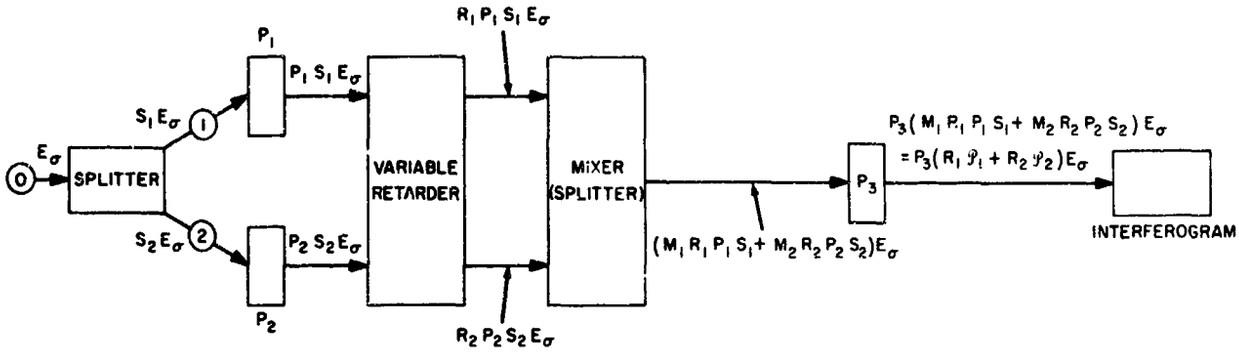


Figure 39-1. Proposed Two-beam Optical Interferometer For Measuring Polarization

mission or reflection and the polarization introduced into the beams. (For an ideal Michelson interferometer $S_1 = S_2 = (1/\sqrt{2})I$, where I is the unit matrix.) The beams 1 and 2 are then passed through two polarizers, P_1 and P_2, respectively. Let P_1 and P_2 represent the Jones matrices of the two polarizers; then, $P_1 S_1 E_\sigma$ and $P_2 S_2 E_\sigma$ will be the amplitudes of the emergent beams. The interferometer now introduces a relative retardation between them which is characterized by the Jones matrices R_1 and R_2. If the relative retardation in phase is $\delta = 2\pi\sigma\tau$ (τ = path difference), one way of expressing R_1 and R_2 is

$$R_1 = e^{-i\delta/2} I, \; R_2 = e^{i\delta/2} I. \tag{39-5}$$

The two beams, which now have the amplitudes $R_1 P_1 S_1 E_\sigma$ and $R_2 P_2 S_2 E_\sigma$, are then recombined either directly (as in Young's experiment) or after a second p e through the splitter (as in Michelson's interferometer). This mixer action can be represented by the Jones matrices M_1 and M_2 for the two beams. Hence, the combined emergent amplitude will be

$$(M_1 R_1 P_1 S_1 + M_2 R_2 P_2 S_2) E_\sigma = (R_1 M_1 P_1 S_1 + R_2 M_2 P_2 S_2) E_\sigma$$

in view of Eq. (39-5). In the radio region, it is possible to deal directly with this combined amplitude in the form of an output voltage of the mixer element of the circuit. On the other hand, in the optical region, it is necessary to introduce an analyzer in the form of a linear polarizer P_3 in the interfering beams. If P_3 is the Jones matrix for the analyzer, the final amplitude will be $E(P_3) = P_3(R_1 M_1 P_1 S_1 + R_2 M_2 P_2 S_2) E_\sigma$, which gives the coherency matrix $J(\delta)$ according to Eq. (39-1). The contribution to intensity for a fixed path difference τ from radiation in the wavenumber interval $d\sigma$ will be $I_\tau(\sigma)\, d\sigma = Tr\, J(\delta)\, d\sigma$. The total intensity is given by

$$I(\tau) = \int_0^\infty I_\tau(\sigma)\, d\sigma; \tag{39-6}$$

the variable part of $I(\tau)$ is known as the interferogram. Although both the variable and constant parts of $I(\tau)$ contain information about the Stokes parameters of the original radiation, only the interferogram can be used for this purpose in the method of Fourier spectroscopy.

39-4 ANALYSIS OF THE INTERFEROGRAM

In order to obtain the Stokes parameters of the incident radiation, it is necessary first to examine how they enter into the observed interferogram.

39-4.1 Computation of $I_\tau(\sigma)$

Letting $\mathcal{P}_j = M_j P_j S_j$, $(j = 1, 2)$, $\mathcal{P}_3 = P_3$, and writing

$$\mathcal{P}_i(\sigma) = \begin{pmatrix} \mathcal{P}_{i1} & \mathcal{P}_{i2} \\ \mathcal{P}_{i3} & \mathcal{P}_{i4} \end{pmatrix}, \; (i = 1, 2, 3), \tag{39-7}$$

and making use of Eq. (39-5) gives the emergent amplitude vector

$$E(\mathcal{P}_3) = \begin{pmatrix} \mathcal{P}_{31} & \mathcal{P}_{32} \\ \mathcal{P}_{33} & \mathcal{P}_{34} \end{pmatrix} \left[\begin{pmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{13} & \mathcal{P}_{14} \end{pmatrix} e^{-i\delta/2} + \begin{pmatrix} \mathcal{P}_{21} & \mathcal{P}_{22} \\ \mathcal{P}_{23} & \mathcal{P}_{24} \end{pmatrix} e^{i\delta/2} \right] \begin{pmatrix} E_x \\ E_y \end{pmatrix},$$

which can be written as

$$\mathcal{E}(\mathcal{P}_3)=\begin{pmatrix}\left(\alpha_1\cos\frac{\delta}{2}-i\beta_1\sin\frac{\delta}{2}\right)E_x \\ \quad+\left(\alpha_2\cos\frac{\delta}{2}-i\beta_2\sin\frac{\delta}{2}\right)E_y \\ \left(\alpha_3\cos\frac{\delta}{2}-i\beta_3\sin\frac{\delta}{2}\right)E_x \\ \quad+\left(\alpha_4\cos\frac{\delta}{2}-i\beta_4\sin\frac{\delta}{2}\right)E_y\end{pmatrix}, \quad (39\text{-}8)$$

where the α's and β's are functions of σ through the $\mathcal{P}_{ij}$'s. Then, computing $J(\delta)$ and taking its trace, gives

$$I_\tau(\sigma)=a_1(\sigma)+a_2(\sigma)\cos\delta+a_3(\sigma)\sin\delta, \quad (39\text{-}9)$$

where, after using Eq. (39-2),

$$a_i(\sigma)=p_i(\sigma)I(\sigma)+q_i(\sigma)Q(\sigma)+r_i(\sigma)U(\sigma)+s_i(\sigma)V(\sigma), \quad (i=1,2,3), \quad (39\text{-}10)$$

and p_i, q_i, r_i, and s_i are known real functions of the $\mathcal{P}_{ij}$'s through the α's and β's. The observed intensity will be $d(\sigma)I_\tau(\sigma)$, where $d(\sigma)$ is a factor representing the instrumental sensitivity.

Introducing the auxiliary quantities $\rho(\sigma)$ and $\psi(\sigma)$ by the equations

$$\left.\begin{aligned}a_2(\sigma)&=\rho(\sigma)\cos\psi(\sigma)\\ a_3(\sigma)&=\rho(\sigma)\sin\psi(\sigma)\end{aligned}\right\}, \quad (39\text{-}11)$$

Eq. (39-9) can also be written in the form

$$I_\tau(\sigma)=d(\sigma)[a_1(\sigma)+\rho(\sigma)\cos\{\delta-\psi(\sigma)\}] \quad (39\text{-}12)$$

At this point it should be noted that the effect of the polarizers P_i has been to introduce the argument function

$$\psi(\sigma)=\tan^{-1}\left[\frac{a_3(\sigma)}{a_2(\sigma)}\right]$$

in the cosine term at the right-hand side of this equation, and to weight the constant and variable parts of the intensity with different functions of σ.

39-4.2 Use of Fourier Spectroscopy

The total intensity is obtained by integrating $I_\tau(\sigma)$ over all wavenumbers. Thus, substituting Eq. (39-12) in Eq. (39-6) gives

$$I(\tau)=\int_0^\infty d(\sigma)\{a_1(\sigma)+\rho(\sigma)\cos[2\pi\sigma\tau-\psi(\sigma)]\}\,d\sigma; \quad (39\text{-}13)$$

and the interferogram is given by

$$\text{Var}\,\{I(\tau)\}=\int_0^\infty d(\sigma)\rho(\sigma)\cos[2\pi\sigma\tau-\psi(\sigma)]\,d\sigma. \quad (39\text{-}14)$$

This expression shows that, owing to the use of the polarizers P_i, the interferogram obtained with the present two-beam interferometer is not a symmetrical function of τ. Hence, the spectral distribution of the incident radiation is not the Fourier cosine transform of the interferogram. Therefore, it is necessary to resort to the full exponential transform for deriving $\rho(\sigma)$ and $\psi(\sigma)$, which in turn would give the four Stokes parameters of the incident radiation. If for negative wavenumbers it is assumed that

$$d(-\sigma)=d(\sigma),\ \rho(-\sigma)=\rho(\sigma),\ \psi(-\sigma)=-\psi(\sigma), \quad (39\text{-}15)$$

Eq. (39-14) becomes

$$\text{Var}\,\{I(\tau)\}=\tfrac{1}{2}\int_{-\infty}^\infty d(\sigma)\rho(\sigma)\exp[-i\psi(\sigma)]e^{-i2\pi\sigma\tau}\,d\sigma. \quad (39\text{-}16)$$

Equation (39-16) is true only for an ideal interferometer. In practice, however, one has to take into account the following imperfections of the system:[3,11]

(1) The finite extent (τ_1, τ_2) of the interferogram; this is indicated by a positive transfer function $T(\tau)$ in front of the integral on the right-hand side of Eq. (39-14). This function is made to vanish outside the interval (τ_1, τ_2).

(2) The zero point of τ may vary with σ, which introduces a further asymmetry in the interferogram; this is taken into account by introducing a phase term $\{-\phi(\sigma)\}$ in the argument of the cosine factor. Amplitude terms which depend on σ only can be included in $d(\sigma)$.

(3) On account of (2), the values of τ_1 and τ_2 may vary with σ, although their difference $(\tau_2-\tau_1)$ may remain constant, thus causing changes in the geometrical parameters of the interferometer. This interaction, which is the result of the finite size of the source, gives rise to another positive amplitude factor $B(\sigma\tau)$ and an additional phase term $\gamma(\sigma\tau)$.

Hence Eq. (39-14) becomes

$$\text{Var}\,\{I(\tau)\} = T(\tau)\int_0^\infty d(\sigma)\rho(\sigma)B(\sigma\tau)\cos\{2\pi\sigma\tau-\psi(\sigma) -\phi(\sigma)+\gamma(\sigma\tau)\}\,d\sigma. \quad (39\text{-}17)$$

This is the correct expression for any recorded interferogram. Now, in addition to Eq. (39-15), if for negative wavenumbers, it is also assumed that

$$B(-\sigma\tau) = B(\sigma\tau),$$
$$\phi(-\sigma) = -\phi(\sigma),\ \gamma(-\sigma\tau) = -\gamma(\sigma\tau), \quad (39\text{-}18)$$

the interferogram could be expressed by

$$\text{Var}\,\{I(\tau)\} = \tfrac{1}{2}T(\tau)\int_{-\infty}^\infty d(\sigma)\rho(\sigma)S^*(\sigma\tau) \exp\{i[\psi(\sigma)+\phi(\sigma)]\}e^{-i2\pi\sigma\tau}\,d\sigma, \quad (39\text{-}19)$$

where

$$S(\sigma\tau) = B(\sigma\tau)\exp\{i\gamma(\sigma\tau)\} = S^*(-\sigma\tau). \quad (39\text{-}20)$$

If the size of the source is sufficiently small, $S^*(\sigma\tau)$ can be replaced by $S^*(\sigma_o\tau)$, where σ_o is the mean wavenumber of the passband. Then, Eq. (39-19) becomes

$$\frac{2\,\text{Var}\,\{I(\tau)\}}{T(\tau)S^*(\sigma_o\tau)} = \int_{-\infty}^\infty d(\sigma)\rho(\sigma) \exp\{i[\psi(\sigma)+\phi(\sigma)]\}e^{-i2\pi\sigma\tau}\,d\sigma. \quad (39\text{-}21)$$

From Eq. (39-2) it is seen that

$$\frac{2\,\text{Var}\,\{I(\tau)\}}{T(\tau)S^*(\sigma_o\tau)}$$

and

$$[d(\sigma)\rho(\sigma)\times\exp\{i[\psi(\sigma)+\phi(\sigma)]\}]$$

form a Fourier pair. Hence, the full (exponential) transform of the interferogram yields the latter function. In Eq. (39-21), the function $T(\tau)$, representing an amplitude factor for the interferogram, is real. It is the Fourier transform of the scanning function. If the latter is apodized, then $T(\tau)$ is a suitably tapered function, which vanishes at the ends of the path difference interval (τ_1, τ_2) in a smooth way. The phase term $\phi(\sigma)$ represents an important source of error, for it contributes, along with $\psi(\sigma)$, to the destruction of the symmetry of the interferogram. The term $S(\sigma\tau)$, known as the source function, includes both amplitude and phase effects. Now, since $T(\tau)$, $d(\sigma)$ and $S^*(\sigma_o\tau)$ are known functions, one can derive $\rho(\sigma)$ and $[\psi(\sigma)+\phi(\sigma)]$. In order to determine the phase error $\phi(\sigma)$, it is necessary to have one interferogram for which $\psi(\sigma)$ is known. Here it is assumed that the interferometer is sufficiently stable during the time of one set of measurements so that $\phi(\sigma)$ is identical for all interferograms. This will have to be tested in actual practice. Of particular interest are the cases where $\psi(\sigma)\equiv 0$; these will be discussed in Section 39-5. For these cases, $\phi(\sigma)$ can be obtained by the well-known methods of Connes[3] and Forman, Steel, and Vanasse;[12] however, this determination of $\phi(\sigma)$ is arbitrary to within an additive term $2\pi\sigma\tau$, which depends on the choice of the origin of the interferogram. If the same origin is taken for all interferograms, this additional term will be identical for all of them.

Now, any interferogram will give the two functions $\rho(\sigma)$ and $\psi(\sigma)+\phi(\sigma)+2\pi\sigma\tau$; and since $\phi(\sigma)+2\pi\sigma\tau$ is known, it is possible to obtain $\psi(\sigma)$, the phase term introduced by the polarizers. The quantities $\rho(\sigma)$ and $\psi(\sigma)$, in turn, yield $a_2(\sigma)$ and $a_3(\sigma)$ according to Eq. (39-11). Then, from Eq. (39-10), it is seen that there are two relations in the four unknown Stokes parameters. Hence, in addition to the interferogram yielding $\phi(\sigma)$, it is necessary to have two other independent interferograms for deriving all the four Stokes parameters.

39-5 PROPOSED INTERFEROMETRIC ARRANGEMENTS

Although the polarizers P_1 and P_2 can be chosen in an arbitrary manner, the present discussion will be limited to the cases where $\mathcal{P}_1$ and $\mathcal{P}_2$ are equivalent to perfect linear polarizers.

The Jones representation for a perfect linear polarizer making an angle θ with the x axis is given by the real matrix

$$\mathcal{P}(\theta)=f\begin{pmatrix}\cos^2\theta & \cos\theta\sin\theta\\ \cos\theta\sin\theta & \sin^2\theta\end{pmatrix}, \quad (39\text{-}22)$$

where $f=f_i$, $(i=1, 2, 3)$, are the transmission factors. Then, putting $\theta=\theta_i$ and carrying out the computations represented by Eqs. (39-8) to (39-10), it is found that $s_1(\sigma)=s_2(\sigma)=p_3(\sigma)=q_3(\sigma)=r_3(\sigma)=0$ in Eq. (39-10). Hence $a_3(\sigma)$ is a function of $V(\sigma)$ alone and $a_2(\sigma)$ contains $I(\sigma)$, $Q(\sigma)$, and $U(\sigma)$ only; in other words, $V(\sigma)$ gets separated from the other three Stokes parameters. Consequently, no two settings of $\mathcal{P}_1$ and $\mathcal{P}_2$ are completely independent of each other. Any one setting gives an interferogram which by the Fourier spectroscopic method yields $V(\sigma)$ and a linear relation between the remaining three Stokes parameters. Hence, three interferograms would be needed to get complete information about the intensity and state of polarization of the incident light. It is possible, however, to use one of them for obtaining $\phi(\sigma)$ by a proper choice of θ_1 and θ_2, thus obviating the necessity of recording a separate interferogram for that purpose.

Table 39-1 shows three simple arrangements: $A_1(\mathcal{P}_1=x, \mathcal{P}_2=x, \mathcal{P}_3=s)$, $A_2(\mathcal{P}_1=y, \mathcal{P}_2=y, \mathcal{P}_3=s)$, and $A_3(\mathcal{P}_1=x, \mathcal{P}_2=y, \mathcal{P}_3=s)$, where x, y, and s stand for linear polarizers with $\theta=0°$, 90°, and 45°, respectively. It is seen that in cases A_1 and A_2, $a_3(\sigma)=0$ and $a_2(\sigma)\neq 0$, which makes $\psi(\sigma)=0$ and $\rho(\sigma)=a_2(\sigma)$, with

$$\rho(\sigma)=\frac{f_1f_2f_3^2}{2}[I(\sigma)+Q(\sigma)]$$

for A_1, and

$$\rho(\sigma)=\frac{f_1f_2f_3^2}{2}[I(\sigma)-Q(\sigma)] \text{ for } A_2.$$

Hence, both interferograms give $\phi(\sigma)$ and the two together yield $I(\sigma)$ and $Q(\sigma)$. In the third case A_3,

$$a_2(\sigma)=\frac{f_1f_2f_3^2}{2}U(\sigma)$$

and

$$a_3(\sigma)=-\frac{f_1f_2f_3^2}{2}V(\sigma);$$

therefore, the Fourier analysis of this interferogram gives these other two Stokes parameters. A test of the stability of the instrument can be made by recording the interferograms in the order A_1, A_3, A_2 and by verifying that $\phi(\sigma)$ is not altered between the first and the last setting.

39-6 CONCLUSIONS

The usual Fourier spectroscopic method is based on the property that when no polarizers are introduced in the interferometer, the intensity of radiation of wavenumber σ at path difference τ is given by $I_\tau(\sigma)=(f_1^2+f_2^2+2f_1f_2\cos\delta)I(\sigma)$, which can be easily obtained by adding the intensities for cases A_1 and A_2 and multiplying by $(2/f_3^2)$ to remove the effect of the analyzer. This is but one possibility, however, which has been utilized for obtaining high-resolution spectra by using the Fourier transformation. This paper has demonstrated the full capability of the interferometric technique in the measurement of both the intensity and state of polarization within any given spectral band. It is hoped that this theoretical discussion will stimulate the workers in the field to build suitable instruments for this purpose.

It may be pointed out that the instrumental imperfections such as imperfect polarizers, dependence of their performance on wavelength, effect of beam splitter, and so forth, have, in principle, been taken into account in this discussion. It should not be impossible, in practice, to take care of these factors

Table 39-1. Interferometric Arrangements With Linear Polarizers

	A_1	A_2	A_3
1 $\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3$	x, x, s	y, y, s	x, y, s
2 Amplitude $E(\mathcal{P}_1) = \mathcal{P}_1 E$	$f_1 E_x \begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$f_1 E_y \begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$f_1 E_x \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
Amplitude $E(\mathcal{P}_2) = \mathcal{P}_2 E$	$f_2 E_x \begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$f_2 E_y \begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$f_2 E_y \begin{pmatrix} 0 \\ 1 \end{pmatrix}$
Amplitude $E(\mathcal{P}_3)$ $= \mathcal{P}_3(R_1\mathcal{P}_1 + R_2\mathcal{P}_2)E$	$\frac{f_3}{2}(f_1 e^{-i\delta/2} + f_2 e^{i\delta/2}) E_x \begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$\frac{f_3}{2}(f_1 e^{-i\delta/2} + f_2 e^{i\delta/2}) E_y \begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$\frac{f_3}{2}(f_1 E_x e^{-i\delta/2} + f_2 E_y e^{i\delta/2}) \begin{pmatrix} 1 \\ 1 \end{pmatrix}$
3 Coherency Matrix $J(\delta)$	$\frac{f_3^2}{4}(f_1^2 + f_2^2 + 2f_1 f_2 \cos\delta) J_{xx} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$\frac{f_3^2}{4}(f_1^2 + f_2^2 + 2f_1 f_2 \cos\delta) J_{yy} \times \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$\frac{f_3^2}{4}[f_1^2 J_{xx} + f_2^2 J_{yy} + f_1 f_2 \{(J_{xy} + J_{yx}) \cos\delta - i(J_{xy} - J_{yx}) \sin\delta\}] \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$
4 Intensity = Tr $J(\delta)$ in terms of the elements of the original coherency matrix J	$\frac{f_3^2}{2}(f_1^2 + f_2^2 + 2f_1 f_2 \cos\delta) J_{xx}$	$\frac{f_3^2}{2}(f_1^2 + f_2^2 + 2f_1 f_2 \cos\delta) J_{yy}$	$\frac{f_3^2}{2}[f_1^2 J_{xx} + f_2^2 J_{yy} + f_1 f_2 \times \{(J_{xy} + J_{yx}) \cos\delta - i(J_{xy} - J_{yx}) \sin\delta\}]$ or $\frac{f_3^2}{2}[f_1^2 J_{xx} + f_2^2 J_{yy} + 2f_1 f_2 \lvert J_{xy}\rvert \cos(\beta_{xy} - \delta)]$
5 Intensity in terms of the Stokes parameters of the incident radiation	$\frac{f_3^2}{4}(f_1^2 + f_2^2 + 2f_1 f_2 \cos\delta)(I + Q)$	$\frac{f_3^2}{4}(f_1^2 + f_2^2 + 2f_1 f_2 \cos\delta)(I - Q)$	$\frac{f_3^2}{4}[(f_1^2 + f_2^2) I + (f_1^2 - f_2^2) Q + 2f_1 f_2 (U \cos\delta + V \sin\delta)]$

because what is involved is the *in situ* determination of $\Phi_i(\sigma)$ by a proper method of calibration. The usual limitations about the source size, bandwidth, frequency of sampling, and so forth, also apply in these polarization measurements.

Finally, it should be mentioned that the proposed method relies heavily on the deliberate introduction of a large asymmetry in the interferogram. Firstly, this requires recording the interferograms on both sides of the origin. Secondly, the process of information retrieval which necessitates the use of exponential transforms is known to be nonlinear. However, this nonlinearity will be of real consequence only when the signal-to-noise ratio is small.

References

1. Fellgett, P. (1951) Thesis, Cambridge University.

2. Jacquinot, P. (1958) *J. Phys. Radium* 19: 39.

3. Connes, J. (1961) *Rev. d'Optique*, 40: 45 116, 171, 231. (Translated into English as U.S. Naval Ordnance Test Station NAVWEPS Report 8099 Nots TP 3157, China Lake, California, 1963).

4. Loewenstein, E. V. (1966) *Appl. Opt.* 5: 845.

5. Wolf, E. (1959) *Nuovo Cimento* (*Ser. X*) 13: 1165.

6. Parrent, G. B., Jr., and Roman, P (1960) 15: 370.

7. Born, M., and Wolf, E. (1959) *Principles of Optics*, Pergamon Press, London, p. 542.

8. Cohen, M. H. (1958) *Proc. IRE* 46: 178.

9. Ko, H. C. (1962) *Proc. IRE* 50: 1950.

10. Ko, H. C. (1967) *IEEE Trans. Antennas and Propagation* AP-15: 10.

11. Steel, W. H. (1967) *Interferometry*, Cambridge University Press, Chap 10.

12. Forman, M. L., Steel, W. H., and Vanasse, G. A. (1966) *J. Opt. Soc. Am.* 56:1, 59.

Discussion

Q. (James L. Lauer, Sun Oil Co.): How sensitive is your method to determine status of polarization, especially when you have no prior knowledge?

A. (A. Fymat, Jet Propulsion Laboratory): If the residual polarization introduced by the beam splitter is properly compensated for, for example by an in-situ calibration, then the sensitivity of the method is essentially that of the more conventional Fourier spectroscopy, for spectral energy distribution. In conventional polarimeters (compare works of Sekera, Dollfus, Gehrels), degrees of linear polarization of about 0.001 percent are obtained. Because of the several advantages of Fourier spectroscopy, a similar sensitivity can at least be expected, with the additional advantages of high resolution (possible resolution of line profiles) and wide spectral range.

In the case of the atmospheres of the major planets, theoretical computations for a semi-infinite Rayleigh scattering model show that degrees of polarization of 80 to 90 percent in the core to about 50 percent in the wings of very weak, weak, strong, and very strong lines are obtained. With the high resolution ($\sim 5 \times 10^{-3}$ cm^{-1}) now made possible by the work of the Connes, such lines can be resolved and the method proposed here can doubtless measure this polarization. I am mentioning this example not as an answer to your question, but as an illustration of the type of phenomena I am interested in, and which are responsible for the conception of the method.

Contents

40. Time Series Analysis in Fourier Spectroscopy*

R. D. Larsen
Illinois Institute of Technology
Chicago, Illinois 60616

Abstract

The formalism and relevance to Fourier spectroscopy of the discipline of statistical spectral analysis of time series is outlined. Four aspects of the symphysis of the two disciplines are discussed: (1) statistical estimation of stochastic spectra, (2) coherence theory, (3) modelling by linear systems, and (4) the relationship of apodization to smoothing and spectral window carpentry.

40–1 INTRODUCTION

Fourier transform spectroscopy (FTS) has experienced an incredible evolution in the past decade, experimentally, theoretically, and computationally. As with most evolutionary processes, Fourier spectroscopy has developed in a natural way, within the environment of its parent; in this case, predominantly within the consanguinity of optical physics. Furthermore, it is natural for related disciplines to develop much like fraternal twins, that is, in parallel, but quite independently of each other. For example, Fourier spectroscopy has much in common with x-ray diffraction, both having optical origins. There are, however, features of the respective data analysis formalisms that are so strikingly similar that each warrants the attention of the other.

Hence, as filiation is socially (if not scientifically and legally) acceptable when the parentage is admitted, the writer would like to direct attention to time series analysis (TSA) and the role that it plays. It is hoped that the genetic advantages to Fourier spectroscopy accruing from time series analysis will become evident throughout the adolescence of Fourier spectroscopy.

This chapter is not intended to be an exposition of the formalism and merits of time series analysis and yet it must be this in part in order to show that time series analysis provides a generalization of the analytic and data analysis formalism of Fourier spectroscopy. The propinquity to transform spectroscopy of a specific diffraction application, diffraction by liquids, is briefly discussed in Appendix 40-A.

Time series analysis is a somewhat little-known subdiscipline of mathematical statistics because, like Fourier transform spectroscopy, it has been developed principally within the past fifteen years. Perhaps the applications of time series analysis to radar signal detection, seismic signal detection, and signal detection and analysis within biomedical engineering

* Supported in part by The National Science Foundation.

(electroencephalographic and electrocardiographic signals) have been largely responsible for its rapid maturation. It is significant that it was within the computational context of time series analysis that the power and expediency of the fast Fourier transform (FFT) was rediscovered.[1] The use of one of the class of fast Fourier transformation algorithms is becoming, and should be, central to any sensible computation of a (Fourier) spectrum. The original suggestion of the use of the factorization property of discrete Fourier coefficients[2] was in the context of an x-ray diffraction (diffraction by liquids), the formalism of which is surprisingly similar (the differences amounting to a change of variables, physical significance, and some other trivial differences) to Fourier transform spectroscopy.[3] It is perhaps worth noting that transform spectroscopy is now considerably ahead of x-ray diffraction in recognizing the utility of Fourier transformation a la fast Fourier transform. However, the next decade must see widespread use of fast transformation techniques not only in the x-ray diffraction field but in all applied Fourier analysis disciplines.

40–2 STATISTICAL SPECTRAL ANALYSIS OF TIME SERIES

40–2.1 Definition of a Time Series

A time series is defined operationally as a set of observations arranged in chronological order. "Time" is a somewhat unfortunate adjective, inasmuch as time need not be the independent variable. For example, space, field, or angle may just as well be the independent variable. In such instances, this paper refers to spatial time series, whereas, if time is the independent variable, it refers to temporal time series. It is convenient to restrict discussion to the temporal time series. Observations may be made or recorded for a set of time points T. A specific measurement at a time t is $X(t)$. A time series is then the set of all observations $[X(t), t \in T]$.

A statistical analysis of a time series is predicated upon associating a specific observation $X(t)$ with a random variable,[4] such that a set of observations $[X(t), t \in T]$ or a time series, is an observation of a stochastic process or ensemble of such observations. A time series is, in this context, considered to be an ordered set of random variables, a given set being a realization of a stochastic process or an ensemble of such sets of random variables.

Further, the sets of all possible time variables T have no restrictions placed on them, but convenient cases are provided if T is a finite discrete set or a finite continuous set: $T = [0, \pm 1, \pm 2, \ldots, N]$ or

$$T = [0 \leq t \leq N].$$

The advantage of the statistical approach to the analysis of time series is that, as in statistical mechanics, the properties of time series, like observations of thermodynamic variables, are expressible as averages over an ensemble of possible observations. The object of a statistical analysis of time series is to ascertain the asymptotic (statistical) properties of a time series from a single observation or realization of the stochastic process.

40–2.2 A Model for a Time Series

As a time series refers to a particular set of measurements, or observations of these, it is of interest to understand how a set of observations is generated by a specific process, that is, its mechanism. Furthermore, it is of interest to be able to employ the predictive power of the mechanism to forecast future time series behavior. Both of these aims are satisfied by synthesizing a model for a time series.

Of most use to Fourier spectroscopy is the class of nonparametric physical (rather than empirical) models for which a given time series is characterized by its autocorrelation function or its spectrum. Spectral analysis is a nonparametric approach. It is often quite convenient, however, to represent a physical situation by a linear system and then to perform a spectral analysis on the input and output of this system. Such an approach is a combination of a modelling and a spectral analysis, which will be discussed in more detail in the next sections.

40–2.3 Statistical Spectral Analysis

40–2.3.1 DEFINITION OF THE SPECTRUM

Statistical spectral analysis is an elucidation of the frequency components of an observed time series. Parzen[5] has suggested that a spectral analysis consist of three distinct aspects: (1) a careful definition of the spectrum of a specific stochastic process, (2) computation of the spectrum by essentially one of four methods: filtering, construction of a smoothed periodogram, covariance averaging, or autoregression, and (3) interpretation of the spectrum, including estimation of the spectral density. The subsequent condensation of the elements of spectral analysis follows Parzen,[5] where the reader can find further details.

An observed time series is neither periodic nor square-summable, so a standard deterministic harmonic analysis or inversion cannot be carried out. A generalized harmonic analysis is, however, possible for a random variable $X(t)$ if there exists an asymptotic correlation function of $X(t)$, $\rho(u)$,

$$\rho(u) = \lim_{T \to \infty} \rho_T(u), \tag{40-1}$$

where the sample correlation function $\rho_T(u)$ is

$$\rho_T(u) = R_T(u)/R_T(0), \quad u = 0, \pm 1, \ldots \tag{40-2}$$

and $R_T(u)$ is the sample convolution function,

$$R_T(u)=\frac{1}{T}\sum_{t=1}^{T-|u|}X(t)X(t+|u|). \quad u=0,\pm1,\ldots \tag{40-3}$$

The spectrum of a stochastic process is defined in terms of a sample spectral density function,

$$f_T(\omega)=\frac{1}{2\pi T}\left|\sum \exp(-it\omega)X(t)\right|^2. \quad -\pi\le\omega\le\pi \tag{40-4}$$

A normalized sample spectral density is then defined by

$$\mathfrak{f}_T(\omega)=f_T(\omega)/R_T(0). \quad -\pi\le\omega\le\pi \tag{40-5}$$

Moreover, the normalized sample spectral density $\mathfrak{f}_T(\omega)$ is just related to the sample correlation function $\rho_T(u)$ by

$$\mathfrak{f}_T(\omega)=\frac{1}{2\pi}\sum_{|u|<T}\exp(-i\omega u)\rho_T(u), \tag{40-6}$$

and

$$\rho_T(u)=\int_{-\pi}^{\pi}\exp(iu\omega)\mathfrak{f}_T(\omega)\,d\omega. \tag{40-7}$$

Finally, the normalized spectral density function $\mathfrak{f}(\omega)$ is related to the asymptotic correlation function,

$$\rho(u)=\int_{-\pi}^{\pi}\exp(iu\omega)\mathfrak{f}(\omega)\,d\omega. \tag{40-8}$$

Filtered spectra are introduced in terms of the sample spectra in order to effect an estimate of the asymptotic functions. A filtered sample spectral density function $f_T^*(\omega)$ is defined in terms of the estimate $\rho_T^*(u)$ where $\rho_T^*(u)=k_T(u)\rho_T(u)$, and the weighting function $k_T(u)$ has the property that

$$\lim_{T\to\infty}k_T(u)=1.$$

The Fourier transform of $\rho_T^*(u)$ is $f_T^*(\omega)$.

$$f_T^*(\omega)=\frac{1}{2\pi}\sum_{|u|<T}\exp(-iu\omega)\rho_T^*(u)$$

$$=\frac{1}{2\pi}\sum_{|u|<T}\exp(-iu\omega)k_T(u)\rho_T(u). \tag{40-9}$$

Therefore,

$$\rho_T(u)=\int_{-\pi}^{\pi}\exp(iu\lambda)f_T(\lambda)\,d\lambda \tag{40-10}$$

and

$$f_T^*(\omega)=\int_{-\pi}^{\pi}f_T(\lambda)K_T(\omega-\lambda)\,d\lambda, \tag{40-11}$$

where the spectral window $K_T(\omega)$ is

$$K_T(\omega)=\frac{1}{2\pi}\sum_{|u|<T}\exp(-iu\omega)k_T(u). \tag{40-12}$$

Thus, the spectrum of a stochastic process is defined only in the asymptotic limit, $T\to\infty$. The sample spectral density and sample correlation functions are estimates of the asymptotic functions. The filtered spectra for various spectral windows all estimate the asymptotic spectra as a spectral average according to various criteria.

40-2.3.2 COMPUTATION OF THE SPECTRUM

One of the central tasks of time series analysis is to obtain the best estimate (in the statistical sense) of the theoretical spectrum. The best estimate is one for which the estimator has minimum bias and minimum variance. The sample spectral density, $f_T(\omega)$, is asymptotically an unbiased, but not consistent, estimator of the theoretical spectral density. A consistent estimator is obtained by a spectral averaging either through filtering or by smoothing. Covariance averaging[5] as a means for obtaining a consistent spectral estimator will be discussed.

Given a time series, by observation of a stochastic process, $X(t)$, $t=1, 2, \ldots, T$, the record length is easily enlarged by adding zeros to the end of the series so that the sample contains 2^n data points. The sample spectrum can then be obtained through use of the fast Fourier transform algorithm,

$$\Gamma\left(k\frac{2\pi}{N}\right)=\sum_{j=1}^{N} X(j)\exp\left(i\frac{2\pi}{N}jk\right)\cdot$$

$$(k=0, 1, \ldots, N-1) \qquad (\quad)$$

The rudimentary *sample spectral density* (*periodogram*) is just the square of the result,

$$f_T\left(k\frac{2\pi}{N}\right)=\frac{1}{2\pi T}\left|\Gamma\left(k\frac{2\pi}{N}\right)\right|^2, \qquad (40\text{-}14)$$

and the filtered or smoothed periodogram is a spectral average of f_T:

$$f_T^*\left(k\frac{2\pi}{N}\right)=\sum_{j=0}^{N-1} K\left(j\frac{2\pi}{N}; k\frac{2\pi}{N}\right) f_T\left(j\frac{2\pi}{N}\right)\cdot \qquad (40\text{-}15)$$

On the other hand, given f_T as above, it can be fast Fourier transformed to obtain the sample covariance, $R_T(u)$, defined for $u=0, 1, \ldots, M$, where M is the truncation point of the record,

$$R_T(u)=\frac{2\pi}{Q}\sum_{k=0}^{Q-} \exp\left(iuk\frac{2\pi}{Q}\right) f_T\left(k\frac{2\pi}{Q}\right)\cdot \qquad (40\text{-}16)$$

The sample correlations are then just normalized covariances

$$\rho_T(u)=R_T(u)/R_T(0). \qquad (40\text{-}17)$$

Moreover, a filtered sample spectral density can be obtained from the $R_T(u)$ by further fast Fourier transformation,

$$f_T^*\left(k\frac{2\pi}{Q}\right)=\frac{1}{2\pi}\sum_{|u|<M}\exp\left(-iuk\frac{2\pi}{Q}\right)k_T(u)R_T(u). \qquad (40\text{-}18)$$

The filtered normalized sample spectral density for $k=0, 1, \ldots, Q$, is then just:

$$\bar{f}_T\left(k\frac{2\pi}{Q}\right)=f_T^*\left(k\frac{2\pi}{Q}\right)/R_T(0). \qquad (40\text{-}19)$$

The method of computing the spectrum via covariances, while extensively used in an earlier era, is perhaps still the best method if the fast Fourier transform algorithm is used to transform the periodogram f_T.[5] The route to the covariance average f_T^* via the fast Fourier transform is thus

$$f_T\left(k\frac{2\pi}{Q}\right)\rightarrow R_T(u)\rightarrow f_T^*\left(k\frac{2\pi}{Q}\right)\cdot \qquad (40\text{-}20)$$

In going from $R_T(u)$ to $f_T^*\left(k\frac{2\pi}{Q}\right)$ the $R_T(u)$ are averaged over a lag window or covariance averaging kernel $k\left(\frac{u}{M}\right)$ which depends on the truncation point $M<T$. Several varieties of covariance averaging kernels have been advocated from time to time.

The Bartlett window (covariance averaging kernel) is

$$k_B(u)=1-|u|. \qquad (40\text{-}21)$$

The Tukey kernel is

$$k_T(u) = \tfrac{1}{2}(1+\cos \pi u) \quad |u|<1$$
$$= 0. \quad |u|>1 \qquad (40\text{-}22)$$

The Parzen window is:

$$k_p(u) = 1-6u^2+6|u|^3 \quad |u|\leq 0.5$$
$$= 2(1-|u|)^3 \quad 0.5\leq|u|\leq 1.0$$
$$= 0. \quad |u|\geq 1 \qquad (40\text{-}23)$$

A modified Parzen window (Parzen II) is

$$k_{PII}(u) = (1+6u^2)k_P(u). \qquad (40\text{-}24)$$

The criterion for choosing between various covariance averaging kernels is a design procedure known as window carpentry. The interested reader is referred to an excellent discussion of this procedure by Jenkins and Watts.[6]

Akaike has recently suggested a criterion for low-pass filter design which incorporates a large variety of covariance kernels.[7]

40–3 COHERENCE THEORY

40–3.1 Definition of the Spectrum

Inasmuch as Fourier spectroscopy derives from optical interferometry, and as it is clear that optical coherence phenomena are manifestations of correlations between fluctuations in the electromagnetic field of a beam of light, it is clear that coherence theory, the statistical theory of optics, must be able to accommodate the basically stochastic phenomenon which is the subject matter of the Aspen conference and of this chapter.

To establish the major premise further, therefore, let us say this in other words. The wave function appropriate to the electromagnetic field of a beam of light is a stochastic quantity and the individual amplitudes of light waves are realizations of an ensemble of such amplitudes. Thus, an electromagnetic waveform is a time series.

Perhaps the central construct of coherence theory[8] the mutual coherence function $\Gamma_{ij}(\tau)$ which is defined as an ensemble average of the analytic signal $V(\mathbf{r}, t)$,

$$\Gamma(\mathbf{r}_i, \mathbf{r}_j, \tau) = \Gamma_{ij}(\tau) = \langle V^*(\mathbf{r}_i, t_1)(V(\mathbf{r}_j, t_2)\rangle. \qquad (40\text{-}25)$$

$V(\mathbf{r}, t)$ is averaged over all realizations of the electromagnetic field, and thus $\Gamma_{ij}(\tau)$ is an ensemble-averaged space-time field correlation function.[9,10] For a stationary stochastic process it is possible to explicitly show the averaging process by

$$\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} V^*(\mathbf{r}_1, t)V(\mathbf{r}_2, t+\tau)\, dt. \qquad (40\text{-}26)$$

$\Gamma_{ij}(\tau)$, being a correlation function, is either a cross-correlation function or an autocorrelation function according to whether $i\neq j$ or $i=j$, respectively; additionally, $\Gamma_{ij}(\tau)$ is itself an analytic signal. The analytic signal $V(\mathbf{r}, t)$ is, however, the manifestation of the stochastic process, and it is therefore a time series (either as a single value or as a digitized or analog record).

The expression $\Gamma(\mathbf{r}, \tau)$, as defined in the asymptotic limit, is a theoretical quantity (a theoretical correlation function) or in time series language, an asymptotic correlation function of the random variable $V(\mathbf{r}, t)$.

The coherence theory literature[11] amply documents the correct definition of the light spectrum. That is, due to the Wiener-Khinchine theorem, the (power) spectrum ($i=j$) or so-called mutual spectral density ($i\neq j$) of light at two field points is

$$W(\mathbf{r}_i, \mathbf{r}_j, \nu) = \int_{-\infty}^{\infty} \Gamma(\mathbf{r}_i, \mathbf{r}_j, \tau) \exp(2\pi i\nu\tau)\, d\tau. \qquad (40\text{-}27)$$

Implicit is the definition of the spectrum in the asymptotic limit $T\to\infty$. Let us refer in this limit to the Fourier transform of the mutual coherence function as *the asymptotic spectrum*.

40–3.2 Estimation of the Mutual Coherence Function

The central point is this. While due obeisance is paid to the basic stochasticity of optical coherence phenomena and while, indeed, the "light spectrum" is correctly defined in the sense of a generalized harmonic analysis, the operational consequences of the sto-

chasticity are virtually ignored in the literature. For the problem is not merely one of transforming the mutual coherence function but, rather, of obtaining the best *estimate* (in the statistical sense) of both the mutual coherence function and the asymptotic spectrum, inasmuch as both are only defined in a limiting sense and are experimentally inaccessible. That which is obtained experimentally is a realization of the stochastic process, however, and is itself an estimate of the mutual coherence function.

The quintessence of a statistical spectral analysis of a time series is, as mentioned in Section 40-2.3.2, the achievement of a *consistent* spectral estimator. The fact has also been mentioned that by the method of covariance averaging, using the fast Fourier transform, such a consistent estimator of the theoretical (asymptotic) spectral density can be obtained.

To put the argument in time series language, the mutual coherence function (for $i=j$) plays the role of a theoretical autocovariance function, the Fourier transform of which defines the spectrum, which is the limit of the first moment of the sample spectrum estimator $C(\nu)$. That is,

$$W(\mathbf{r}_i, \mathbf{r}_i, \nu) = \lim_{T\to\infty} E[C(\mathbf{r}_i, \mathbf{r}_i, \nu)] = \int_{-\infty}^{\infty} \Gamma(\mathbf{r}_i, \mathbf{r}_i, \tau) \exp(-2\pi i\nu t)\, dt. \quad (40\text{-}28)$$

Further, by the Wiener-Khinchine theorem

$$\Gamma(\mathbf{r}_i, \mathbf{r}_i, \tau) = \int_{-\infty}^{\infty} W(\mathbf{r}_i, \mathbf{r}_i, \nu) \exp(2\pi i\nu t)\, d\nu. \quad (40\text{-}29)$$

For finite T, however,

$$E[C(\mathbf{r}, \nu)] = \int_{-T}^{T} \Gamma(\mathbf{r}, t)\left[1-\frac{t}{T}\right]\exp(-2\pi i\nu t)\, dt.$$

Thus $E[C(\mathbf{r}, \nu)]$ is biased for finite T but is an asymptotically unbiased estimator as the record length T tends to infinity. It can be shown, however, that the biased estimator has the smallest mean square error.[6]

40-4 MODELLING BY LINEAR SYSTEMS

40-4.1 Linear Stochastic Filtering

Let us refer to a full or a complete time series analysis as one which includes a statistical spectral analysis in conjunction with a modelling of the underlying stochastic process by a linear system. Of course, recent advances have been made in nonlinear filtering, but that subject will be neglected here.

Classical filtering theory calls for either the determination of the output of a linear system (filter) given the input and the response function of the filter, or an analysis of the system given the input and output of observations.

However, a standard (deterministic) harmonic analysis and inversion employing a given filter spectral response function is, of course, not applicable to stochastic processes. The generalization is provided by a linear stochastic filtering. This problem was classically handled by Wiener but his method has since been somewhat superseded by use of the so-called Kalman-Bucy filter, which involves a solution, not of the Wiener-Hopf equation, but of the Ricatti equation.[12,13,14] The Kalman-Bucy technique involves characterizing a stochastic process in terms of a Markov process rather than through a covariance function, and the corresponding optimal filter involves a solution of a set of differential equations rather than the determination of an impulse response function. Despite several successes achieved by the Kalman-Bucy technique (and the equivalence of the results for the linear theory) the Wiener-type spectral methods are, perhaps, more easily visualized, because they are more algebraic. Furthermore, the computation of covariance averages using the fast Fourier transform should swing the pendulum of fashion back somewhat toward the covariance approach.

It is in terms of a linear time-invariant (stochastic) filter that the complete generality of the notion of a stochastic spectrum can best be appraised.[5] The writer's viewpoint is that Fourier spectroscopy and, for example, liquid diffraction can be considered to be special cases of linear stochastic filtering coupled with spectral analysis of the physical time series embedded within the respective disciplines.

The literature amply documents the treatment of an interferometer as a linear system[15,16] with a concomitant determination of the impulse response of the interferometer. The point here is simply to suggest that any linear system modelling should involve a linear stochastic filtering for reasons specified in Sections 40-2 and 40-3.

For linear stochastic systems the output autocorrelations can be written in terms of input autocorrelations through the use of stochastic Green's functions. Section 40-5 shows the relationship of such general stochastic kernels to the apodization functions of Fourier spectroscopy.

40-5 STOCHASTIC GREEN'S FUNCTIONS AND APODIZATION

40-5.1 The Spectral Density Green's Function

The author would like to discuss what he believes to be the generalization of apodization in Fourier spectroscopy in terms of time series analysis, namely, that apodization is a special case of window carpentry modulo time series analysis.

The impulse response function, which has been referred to, and which characterizes a linear system is, of course, known in Fourier spectroscopy as a scanning function. A linear system influences or modifies an input by broadening, smearing, spreading, or attenuating it. Thus, the ubiquitous system response function is called, in diffraction theory, the modification function (frequency domain), and peak shape function (time domain). It is also called variously in other disciplines, the apparatus or instrument function, spread function, and slit function. In spectral analysis one refers to lag or data (time domain) windows and spectral (frequency domain) windows.

Apodization, the process by which the shape of a scanning function is adjusted so as to attenuate the side lobes, is, in time series language, the process of designing a lag window or covariance averaging kernel so as to increase its focusing power.

Let us consider Fourier spectroscopy to be a specific physical model of a stochastic process (see Section 40–3 for the nature of the stochasticity) such that by means of a linear stochastic filtering, a stochastic spectrum (asymptotic spectral density as characterized in Section 40–2) is obtained. Then, the general form of the resulting spectral density can be considered to be the result of filtering by a stochastic kernel or stochastic Green's function.

The basic idea is that instead of solving for y as a function of x (using standard differential operator notation[3]) in

$$Ly=x(t), \tag{40-30}$$

in terms of a Green's function $G(t, \tau)$, the problem is one of solving for a spectral density Φ_y, in terms of an input function Φ_x, and a stochastic Green's function, $K(t, \tau)$,

$$K\Phi_y=\Phi_x(t). \tag{40-31}$$

If H is some linear operation on an input process x, one may consider H to be an ensemble containing operations H_i, each of which has associated with it a Green's function $h_i(t, \tau)$, representing the response of H_i to a unit impulse $\delta(t-\tau)$, at time τ.[18] Thus, the response of H_i to an input $x(t)$ is

$$H_i x(t)=\int_{-\infty}^{\infty} h_i(t, \tau)x(\tau)\, d\tau. \tag{40-32}$$

The general result can be written operationally as

$$\Phi_{Hx}=K_H \cdot \Phi_x, \tag{40-33}$$

or,

$$\Phi_{Hx}(f)=\int_{-\infty}^{\infty} dsK_H(s, f)\Phi_x(s), \tag{40-34}$$

where the stochastic Green's function K_H is

$$K_H(s, f)=\int_{-\infty}^{\infty} d\tau_1 \int_{-\infty}^{\infty} d\tau_2 \int_{-\infty}^{\infty} d\tau \exp(2\pi i\tau f) \exp(-2\pi i(\tau_2-\tau_1)s) \times\langle h(t; \tau_1)h^*(t+\tau; \tau_2)\rangle. \tag{40-35}$$

As a stochastic Green's function depends on the statistical measure, the Green's function K_H is a spectral density Green's function. Thus, if the measure is the spectrum Φ, then

$$\Phi_y(f)=\int K(s, f)\Phi_x(s)\, ds, \tag{40-36}$$

whereas, if the measure is the autocorrelation R,

$$R_y(\beta)=\int G(\beta, \sigma)R_x(\sigma)\, d\sigma. \tag{40-37}$$

Hence, in terms of a general statistical measure μ,

$$\mu(y)=\int K(y, x)\, d\mu(x). \tag{40-38}$$

The general kernel $K(y, x)$ is said to map from the space of x to the space of y and depends on the specific measure μ, chosen to characterize the random variable y.

As far as Fourier spectroscopy is concerned, one is interested in estimating the spectral density, given the interferogram which is itself an experimental estimate of the asymptotic autocovariance function. An estimation of the spectral density is predicated upon a mapping from the space of $B_x(s)$ to the space $B_y(\sigma)$, where $B_x(s)$ is the (power) spectral density of the analytic signal $V(r, t)$. Implicit is the specific choice of the spectral density measure by means of which a stochastic (spectral density) Green's function effects the mapping. This heuristic analysis is thus a generalization of the following equations[19] in Fourier spectroscopy and in conventional spectroscopy, respectively:

$$B'_e(\sigma) = \tilde{A}(\sigma) * \tilde{F}(\sigma) \tag{40-39}$$

and

$$B'(\sigma) = S(\sigma) * B(\sigma). \tag{40-40}$$

40–5.2 Window Carpentry vs Apodization

The impossibility of realizing experimentally a theoretical sample spectral density or sample autocovariance function, let alone the asymptotic functions referred to in Section 40–2, is not unique to Fourier spectroscopy. In Fourier spectroscopy, of course, a truncated interferogram is all that is ever realized. Furthermore, the interferogram must be sampled in order to compute a Fourier transform. Both truncation and sampling have to be carefully effected in transforming to obtain the spectral distribution. However, sampling according to the prescription required by the fast Fourier transform algorithm allows for a completely faithful and reproducible Fourier transform. One can transform back and forth between $F(x)$ and $B_e(\sigma)$ completely promiscuously if the sampling is done correctly with what is referred to as the Lanczos algorithm. The writer has previously[20] commented on this aspect of the Lanczos algorithm, and he feels that the faithfulness of the transformation is as important, if not more important, than the "fastness." For example, covariance averaging as discussed in Section 40–2 involves repeated transformation, so that this method of spectral averaging using the Lanczos algorithm depends on the accuracy with which it can be performed. Thus, it is the band-limited nature of the spectrum $B_e(\sigma)$, that will be discussed further.

Of course, because there is a truncated interferogram, Fourier transformation introduces "wings," "side lobes," "ripples," or "false detail," into the spectral distribution. Apodization seeks to suppress these side lobes by multiplying the interferogram by an apodizing function $A(X)$, before Fourier transformation. In general,

$$B'_e(\sigma) = \int_{-\infty}^{\infty} F(x)A(x) \cos 2\pi\sigma x \, dx. \tag{40-41}$$

One may also obtain an apodized spectrum by convolution of the unapodized spectrum with the Fourier transform of the apodizing function, the scanning function.[21] Let us refer to the result of apodization as a focusing of the spectrum. Various apodizing functions thus have better focusing power than others.

In time series analysis one is interested in obtaining the best estimate to the asymptotic spectrum, that is, spectral estimation. This paper has suggested spectral estimation by means of covariance averaging where the spectral averaging is effected variously by covariance averaging kernels, some of which were outlined in Section 40–2. The Fourier transforms of the covariance averaging kernels or lag windows are spectral windows. Comments have been made on the nature of various spectral windows[3,22,23] and no further comment will be made here. The interested reader is referred to the excellent exposition by Jenkins and Watts.[6]

An optimal spectral window is one that balances the fidelity and stability or bias and variance of the window (or filter). Achieving this balance according to the specific optimality criterion needed in any one situation is the design procedure known as window carpentry. What is desired and achieved in most cases is a spectral window with small side lobes. Spectral averaging with such a window $W(X)$ then generates a mean smoothed spectrum Γ

$$\Gamma(f) = \int_{-\infty}^{\infty} W(g)\Gamma(f-g) \, dg, \tag{40-42}$$

where

$$\Gamma(f) = E[\bar{C}(f)], \tag{40-43}$$

and

$$\bar{C}(f) = \int_{-\infty}^{\infty} k(u)c(u) \exp(-2\pi i f u) \, du, \tag{40-44}$$

where $\bar{C}(f)$ is a smoothed spectral estimator defined in terms of the covariance kernels $k(u)$ and the covariance estimators $c(u)$. Thus, covariance averaging using lag windows with optimum focusing is a generalization to stochastic processes of apodization as commonly formulated for deterministic processes.

References

1. Cooley, J. W., and Tukey, J. W. (1965) *Math. Comp.* **19**: 296.

2. Danielson, G. C., and Lanczos, C. (1942) *J. Franklin Inst.* **233**: 365, 435.

3. Larsen, R. D. (1970) Spectral analysis of time series and coherence theory, *Bull. Am. Phys. Soc.* **15**: 34.

4. Parzen, E. (1967) *Time Series Analysis Papers*, Holden-Day.

5. Parzen, E. (1968) *Statistical Spectral Analysis (Single Channel Case) in 1968*, Technical Report No. 11 (ONR), Department of Statistics, Stanford University.

6. Jenkins, G. M., and Watts, D. G. (1968) *Spectral Analysis and Its Applications*, Holden-Day.

7. Akaike, H. (1967) *Low Pass Filter Design*, Technical Report No. 20 (ARO), Department of Statistics, Stanford University.

8. Born, M., and Wolf, E. (1965) *Principles of Optics*, Pergamon Press, 3rd Ed.

9. Francon, M. (1966) *Optical Interferometry*, Academic Press.

10. Klauder, J. R., and Sudarshan, E. C. G. (1968) Fundamentals of Quantum Optics, Benjamin.

11. Mandel, L., and Wolf, E. (1965) *Rev. Mod. Phys.* **37**: 231.

12. Kalman, R. E. (1960) *J. Basic Engr.* (Trans. ASME) 82D: 34–45.

13. Kalman, R. E., in Proc. Sym. on System Theory, New York, 1965, Polytechnic Press, p. 197.

14. Kalman, R. E., Falb, P. L., and Arbib, M. A. *Topics in Mathematical System Theory*, McGraw-Hill, 1969.

15. Vanasse, G. A., and Sakai, H. (1967) Fourier spectroscopy, in *Progress in Optics*, Vol. VI, North-Holland Publishing Co., p. 264.

16. Stoner, J. O. (1966) *J. Opt. Soc. Am.* **56**: 371.

17. Lanczos, C. (1961) *Linear Differential Operators*, Van Nostrand.

18. Adomian, G. (1964) Stochastic Green's functions, in *Proc. Symposia in Applied Mathematics*, Vol. XVI, *Stochastic Processes in Mathematical Physics and Engineering*, Am. Math. Soc.

19. Vanasse, G. A., and Sakai, H. (1967) Fourier spectroscopy, in *Progress in Optics*, Vol. VI, North-Holland Publishing Co., p. 274.

20. Starshak, A. J., and Larsen, R. D. (1970) *Phys. and Chem. of Liquids* **1**: 303.

21. Vanasse, G. A., and Sakai, H. (1967) Fourier spectroscopy, in *Progress in Optics*, Vol. VI, North-Holland Publishing Co., p. 287.

22. Larsen, R. D., and Starshak, A. J. (1969) *Spectral Analysis of Time Series: Linear Stochastic Filtering in Chemical Physics*, Computer Applications in Chemistry Workshop Text.

23. Larsen, R. D., and Starshak, A. J. (1969) Single channel spectral analysis and spectrum enhancement, 158th National Meeting of The American Chemical Society, Sept. 1969, New York.

24. Starshak, A. J. (1970) Doctoral dissertation, Illinois Institute of Technology.

25. Waser, J., and Schomaker, V. (1953) *Rev. Mod. Phys.* **25**: 671.

Discussion

Comment (R. D. Larsen): Had I known in advance of this conference that the advantages of 10^6 samples were to be a preoccupation of this conference ("happiness is 10^6 samples"), I would have brought slides showing how a similar situation obtains in liquid diffraction. It is well documented in liquid diffraction literature that by collecting data over larger and larger scattering angles, the radial distribution function becomes increasingly better resolved. That is, the larger R_{max} is, like the greater σ_{max} is, the better is the resulting $D(r)$ and $B(\sigma)$, respectively. In liquid diffraction the quest is for the asymptotic radial distribution function (derived from $R_{max} = \infty$), inasmuch as the theoretical radial distribution function of statistical mechanics is this quantity.

Q. (Alain L. Fymat, Jet Propulsion Laboratory): The mutual coherence function (mcf) is only one of the four elements of the mutual coherence matrix (mcm) which, incidentally, is not defined in the optical literature (Born and Wolf, *Principles of Optics;* Beran and Parrent, *Theory of Partial Coherence*). For a single space-time point, this matrix reduces to the polarization coherence matrix (pcm) (also not defined). The Fourier transform of (mcm) for zero-vector spacing yields Fourier analysis; pcm can also be related to this transform.

Because of these various relationships, I wondered whether you have considered extending your approach to also describe the state of polarization of the radiation field. (For your complete information I might add that Professor S. Veno from the University of Tokyo and Professor V. Sobolev from the U.S.S.R. have formulated radiative transfer theory using stochastic theory. Their theory does not, however, consider polarization.)

A. (R. D. Larsen): A unified treatment of second- and higher-order coherence phenomena and the state of polarization is provided by the introduction of higher-order moments of the complex analytic wave amplitude. These general moments were introduced by Wolf and Mandel,[11] and Klauder and Sudarshan.[10] While higher-order correlation effects have been measured, for the simple interference experiment it is usually sufficient to consider the second-order coherence tensor instead of the mutual coherence function. It is interesting to note that whereas we have considered some elements of a generalized harmonic analysis as applied to Fourier spectroscopy, it was Wiener who not only first introduced such a generalization but also first introduced the idea of coherence matrices for describing partial polarization in terms of the degree of polarization P.

More germane to Fourier spectroscopy, I might suggest that the time series approach is to the deterministic (Bracewellian) approach as statistical thermodynamics is to thermodynamics, in that both the latter and deterministic Fourier analysis neglect fluctuations and correlations thereof.

Appendix 40-A

Relation Between Formalisms of Diffraction by Liquids and Fourier Spectroscopy

It is both instructive and amusing to consider the closely parallel features of Fourier spectroscopy and the formalism of x-ray diffraction by liquids. Each technique, of course, provides quantitative structural information (of a different kind in each case) about a substance bathed in radiation from widely separated regions of the electromagnetic spectrum.

One obtains an intensity function, $\langle I(k)\rangle$, in liquid diffraction, where k is proportional to the scattering angle and is a function of frequency,

$$k=\frac{4\pi}{\lambda}\sin\theta, \tag{40-A1}$$

that is, $\langle I(k)\rangle$ is a spectrum and is obtained directly. Furthermore, $\langle I(k)\rangle$ is the Fourier transform of the autocorrelation of the electron density,

$$\langle I(\mathbf{k},t)\rangle=\int\cdot\int\langle\rho(\mathbf{r}_1,t)\rho(\mathbf{r}_2,t)\rangle \exp\left[i\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)\right]d\mathbf{r}_1\,d\mathbf{r}_2. \tag{40-A2}$$

The object of the analysis is to obtain the radial distribution function, $D(r)$, which gives the distribution with distance of individual pair interactions within a liquid $D_{ij}(r_{ij})$,

$$D(r)=\sum_{i<j}D_{ij}(r_{ij}). \tag{40-A3}$$

From the physical information embedded within $D(r)$ it is possible to determine the time-averaged geometrical structure of a liquid. Now, $D(r)$ is related to $I(k)$ by Fourier transformation. As both $D(r)$ and $I(k)$ are odd functions, using reduced variables, one can write

$$D(r)=2\int_0^\infty I(k)\sin(2\pi rk)\,dk. \tag{40-A4}$$

As the scattering of x-rays by a liquid may also be considered to be a stochastic process, the techniques of time series analysis are applicable to liquid diffraction.[24] The problem is one of estimating the autocovariance $D(r)$.

One obtains experimentally a band-limited spectrum, the Fourier transform of which, due to simple truncation, contains side lobes which inhibit a clear-cut structural analysis. Thus, the intensity function is commonly multiplied by a so-called modification function $M(k)$, prior to Fourier transformation.

$$D'(r)=2\int_0^\infty I(k).M(k)\sin(2\pi rk)\,dk. \tag{40-A5}$$

The modification functions are clearly the same functions as the apodizing functions of Fourier spectroscopy, which, if generalized in terms of time series analysis, are covariance averaging kernels or lag windows (more generally, in the frequency domain, spectral density Green's functions); that is, compare Eq. (40-A5) with

$$B'(\sigma)=2\int_0^\infty F(x)A(x)\cos(2\pi\sigma x)\,dx. \quad \text{(40-A6)}$$

The difference between Eq. (40-A5) and Eq. (40-A6) is simply due to the fact that $F(x)$ and $B(\sigma)$ are even functions, and so they are related by a cosine Fourier transformation.

Liquid diffraction does not, however, enjoy the splendid Fellgett advantage, the spectrum $I(k)$ being obtained in a "dispersive" manner by step-scanning through values of k and counting incident (scattered) photons for sizable lengths of time for each k value. Use can be made, however, of the Lanczos algorithm (fast Fourier transform) in liquid diffraction as well. Furthermore, it should be noted that the one-dimensional radial distribution function of liquid diffraction is formally a lower-dimensional analog of the three-dimensional Patterson function of x-ray crystallography. In fact, the use of modification functions in liquid diffraction originated in crystallography.[25]

Contents

41. Increased Spectral Resolution From Fixed Length Interferograms

Alvin M. Despain and Jeffery W. Bell
Utah State University
Logan, Utah 84321

Abstract

Problems of enhancing the spectral resolution of fixed length interferograms are discussed and several methods of implementation are considered. Prolate spheroidal wavefunctions and the method of moments are found to be useful although in many cases signal-to-noise ratio must be drastically decreased to obtain a significant increase in resolution.

41–1 INTRODUCTION

41–1.1 Fourier Spectroscopy

Fourier spectroscopy is the determination of the power spectral density function of radiation with the aid of an interferometer. Michelson[1] invented the two-beam interferometer that bears his name (Figure 41-1) and used it for several purposes, including the determination of the structure of certain optical spectral features. Later, many workers modified, refined, and improved the interferometric technique.[2,3,4] Today Fourier spectroscopy is widely used[5] where its advantages of multiplex detection and large optical throughput outweigh its mechanical problems and the need for Fourier transforming the output signal to obtain spectral data.

The ideal Michelson interferometer, as shown in Figure 41-1, is a two-beam device in which the incident radiation is split into two beams, recombined, and detected. The path traveled by the radiation in one beam is of different length from that traveled by the other and is a function of the mirror position. The impulse response of the device[6] from the entrance aperture to the detector aperture is thus

$$h(t) = \frac{\delta(t)}{4} + \frac{\delta(t - x/c)}{4}, \tag{41-1}$$

where $\delta(t)$ is the impulse, x is the path difference, and c is the velocity of light.

The spectral response to an input spectrum $F(\nu)$ is then the product of the Fourier transform of Eq. (41-1) and $F(\nu)$,

$$G(\nu, x) = \frac{F(\nu)}{4}\left[1 + e^{-j2\pi\nu x/c}\right], \tag{41-2}$$

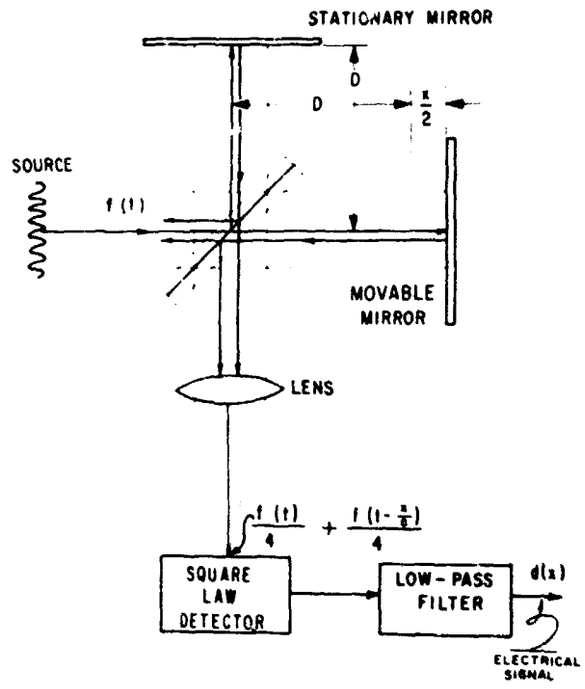


Figure 41-1. Michelson Interferometer Spectrometer

where $\nu = f/c$ (wavenumber). If the detector is assumed to be an ideal square law detector, then its response is

$$d(x) = K_d \int_{\nu_1}^{\nu_2} |F(\nu)|^2 [1 + \cos 2\pi\nu x]\, d\nu, \tag{41-3}$$

where the radiation is limited to the spectral region between ν_1 and ν_2. The signal $d(x)$ is generally recorded (this record is called an interferogram), digitized, and Fourier transformed on a ditigal computer. Thus, the power spectrum of $|F(\nu)|^2$ is recovered by calculating

$$|F(\nu)|^2 = \int_{-\infty}^{+\infty} \left[d(x) - \frac{d(0)}{2} \right] \cos 2\pi\nu x\, dx. \tag{41-4}$$

41-1.2 Spectral Resolution

Note that it is physically impossible to record $d(x)$ for all possible values of x. Therefore, $|F(\nu)|^2$ can only be approximately determined as

$$|\hat{F}(\nu)|^2 = \int_{-x_0}^{x_0} \left[d(x) - \frac{d(0)}{2} \right] \cos 2\pi\nu x\, dx \tag{41-5}$$

$$= \int_{-\infty}^{+\infty} a(x) \left[d(x) - \frac{d(0)}{2} \right] \cos 2\pi\nu x\, dx, \tag{41-6}$$

where

$$a(x) = \text{rect}\left(\frac{x}{2x_0}\right) = \begin{cases} 0, & |x| > x_0 \\ 1, & |x| < x_0. \end{cases} \tag{41-7}$$

Thus, the approximate spectrum is the convolution of $|F(\nu)|^2$ and the transform of $a(x)$, or

$$|\hat{F}(\nu)|^2 = |F(\nu)|^2 * A(\nu), \tag{41-8}$$

where

$$A(\nu) = 2x_0 \,\text{sinc}\,[2\nu x]. \tag{41-9}$$

The sinc function has its first zero at

$$\nu = \frac{1}{2x_0}, \tag{41-10}$$

and by the Raleigh criterion[1] the spectral resolution of $|\hat{F}(\nu)|^2$ is just

$$\delta\nu = \frac{1}{2x_0}. \tag{41-11}$$

Here, then, is one of the major problems of Fourier spectroscopy. How can better spectral resolution be obtained if x_0 cannot be further increased due to physical limitations?

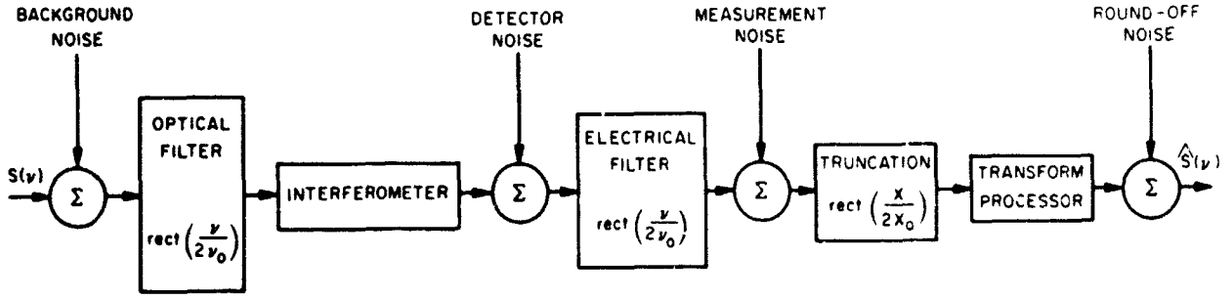


Figure 41-2. Fourier Spectroscopy System

41-1.3 Noise

Resolution is only one measure of the quality of a spectral estimate. The other major factor is the reliability of the result. This is usually described by the variance σ^2 of a sample of the spectral estimate. This variance is caused by noise that enters the system at various points. Signal-to-noise ratio (SNR) is defined as the ratio of the average signal power to the average noise power or

$$\mathrm{SNR} = \frac{\overline{s^2}}{\overline{\sigma_n^2}}, \qquad (41\text{-}12)$$

where

$\overline{s^2}$ = average signal power,

$\overline{\sigma_n^2}$ = average noise power or variance.

Four basic types of noise are important in a Fourier spectroscopy system (Figure 41-2). These are sometimes identified as:

(1) background noise,

(2) detector noise,

(3) measurement noise,

(4) round-off noise.

Background noise is due to unwanted radiation entering an interferometer along with the radiation to be examined. Both signals are assumed to be limited to the same spectral band by an optical bandpass filter or limited detector response. If this is not the case, energy outside the spectral band of interest can degrade the final spectral estimate if the detector noise is increased. The in-band noise remains indistinguishable from the signal.

Detector noise is that noise generated by the detection and amplification process. This noise will generally contain energy that does not correspond to the optical bandpass mentioned above. This energy can usually be rejected by an electrical filter. Again, the in-band noise is indistinguishable from the signal. Thus, the signal, background noise, and detector noise are bandlimited and then truncated.

Measurement noise is due to measurement inaccuracies, generally analog-to-digital quantization errors. This noise is not bandlimited, but since it operates only on a truncated signal, it is truncated.

Round-off noise is due to the round-off errors made in calculating a spectral estimate from the interferogram. In general this noise is neither bandlimited nor truncated.

Attempts to increase the resolution of a spectral estimate must consider the manner in which each of these noises degrades the spectral estimate, if the result is to have any meaning.

41-2 INCREASING RESOLUTION

41-2.1 Information Content

It can be seen from the preceding discussion that an interferometer may be considered as a noisy bandlimited communication channel through which information about the input radiation passes. As such, it will have a finite channel capacity [1]. Thus, a finite length interferogram will contain only a finite amount of information. This information is the product of the total number of orthogonal waveforms or dimensions that can pass through the truncated channel and the amount of information each waveform can carry. The limit to this information is given by the well-known channel capacity theorem to be:

$$C = \log_2 (1 + \mathrm{SNR}) \text{ bits/second} \qquad (41\text{-}13)$$

where the input spectrum is restricted to the interval

$[-\nu_0, \nu_0]$. In the present case, retardation is the independent variable in the interferogram, hence the total information is bounded by

$$I_t = 2\nu_0 x_0 \log_2 (1+\text{SNR}) \text{ bits}, \qquad (41\text{-}14)$$

where $2x_0$ is the length of the interferogram.

41-2.2 Trading SNR for Resolution

Any subsequent processing such as finite Fourier transforming will only destroy some of the above information. Further, note that Eq. (41-14) is not only a measure of the information, but a simple formula relating resolution and SNR, since by Eq. (41-11) the spectral resolution $\delta\nu$ is

$$\delta\nu = \frac{1}{2x_0}. \qquad (41\text{-}15)$$

Now if the interferogram is processed to provide greater resolution, $\delta\nu'$, the maximum possible value is given by the relation

$$\frac{1}{\delta\nu'} \log_2 (1+\text{SNR}') = \frac{1}{\delta\nu} \log_2 (1+\text{SNR}), \qquad (41\text{-}16)$$

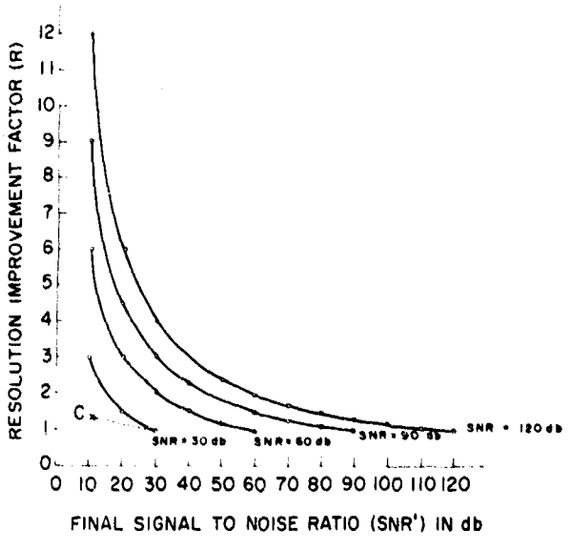


Figure 41-3. Theoretical Signal-to-Noise and Resolution Relationship

where SNR′ is the signal-to-noise ratio in the resulting spectral estimate. Thus, the maximum possible improvement in resolution is

$$R = \frac{\delta\nu}{\delta\nu'} = \frac{\log_2 (1+\text{SNR})}{\log_2 (1+\text{SNR}')}. \qquad (41\text{-}17)$$

This relationship determines the possible improvement in resolution, given the interferogram SNR and the resulting spectral SNR′. This function is plotted in Figure 41-3. It can be readily observed that a very large SNR is required in order to achieve any significant gain in resolution.

41-2.3 Maximum Resolution

The usual finite Fourier transform destroys, to some extent, the information in the interferogram;[10] that is, it is a nonoptimum processor to derive spectral information from an interferogram. Woodward[10] has shown that prolate spheroidal wavefunctions (PSWF) are superior and can thus give better resolution for the same SNR. This is only true for small retardation-bandwidth products $\nu_0 x_0$ since the two methods converge for large $\nu_0 x_0$.[10]

Intuitively, it appears that any *a priori* information about the characteristics of the spectral distribution function could be exploited to improve resolution. Such will be shown to be the case later in this paper.

In general, resolution can in principle be improved in these ways:

(1) trade SNR for resolution,

(2) use optimum transform techniques,

(3) utilize all *a priori* information known about the spectrum under study.

The remainder of this paper concerns the implementation of these three strategies.

41-3 METHODS OF IMPLEMENTATION

The preceding discussion provides some insight into the possibilities of resolution enahancement. Particular techniques for accomplishing this are discussed in the following paragraphs.

41-3.1 Interferogram Extrapolation

If the signal that produced the interferogram is frequency band-limited (as all real radiation must be) and noiseless, then the interferogram $d(x)$ is infinitely differentiable everywhere and a Taylor series expansion derived from a limited portion of $d(x)$ would converge to $d(x)$ everywhere. Thus, a limited length interferogram could be extrapolated, and the resulting spectral resolution would be improved. In practice, the required Taylor series coefficients are difficult to determine. Also, noise exists in any physical system, and this precludes any accurate knowledge of the

higher order coefficients. The authors are unaware of any successful attempts to apply this technique to Fourier spectroscopy.

Another method of extrapolation was suggested by J. L. Harris[11] as a way to improve the resolution of an imaging system. The essence of the method is to estimate the interferogram outside of $|x_0|$ by relating samples of the known interferogram in a system of linear equations to samples outside of the known interferogram. These simultaneous equations are then solved for the values beyond $|x_0|$. The total set of known and extrapolated samples are then Fourier transformed and greater resolution results.

According to Forman[12] J. Connes applied this technique to Fourier spectroscopy and obtained a resolution improvement of 30 percent for a signal-to-noise ratio of 1000:1. Although an assumption had to be made on the final SNR, an attempt was made to plot the results of the Connes findings on Figure 41-3. As might be expected, it appears that this method falls short of the maximum possible improvement and is not of much use in improving the resolution.

41–3.2 Prolate Spheroidal Wavefunctions

A different approach to extrapolation was introduced into the problem of optical restoration by Barnes.[13] His method was based on expanding the signal in terms of prolate spheroidal wavefunctions,[14] several examples of which are illustrated in Figure 41-4. Williams and Chang[15] then applied the technique to the problem of Fourier spectroscopy. Several other investigators[10,16–18] have since continued the work with PSWF.

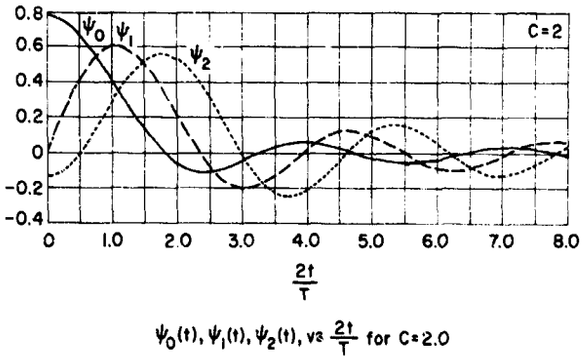


Figure 41-4. Examples of Prolate Spheroidal Wavefunctions[19]

Slepian and Pollak[19] pointed out the usefulness of the PSWF as expansion functions for signal processing applications. For example, the PSWF are orthogonal over both a finite interval $[-x_0 \leq x \leq x_0]$ and the infinite interval $[-\infty < x < +\infty]$. The Fourier transform of a PSWF is a scaled and truncated version of itself. Thus, it is quite convenient to obtain a Fourier transform of an arbitrary bandlimited function, expanded in terms of these functions over a finite interval. Following the notation of Rushforth and Harris[18] for a frequency bandlimited function $g(x)$,

$$g(x) = \sum_{n=0}^{\infty} a_n \psi_n(x), \tag{41-18}$$

where

$$a_n = \lambda_n^{-1} \int_{-x_0}^{+x_0} g(x)\psi_n(x)\, dx, \tag{41-19}$$

and $\psi_n(x)$ is the nth PSWF with $c = 2\pi\nu_0 x_0$.

The self-transform property[19] then allows the calculation of the spectral estimate $G(\nu)$ as

$$G(\nu) = \sum_{n=0}^{\infty} \frac{a_n}{\alpha_n} \operatorname{rect}\left(\frac{\nu}{2\nu_0}\right) \psi_n\left(\frac{x_0\nu}{\nu_0}\right), \tag{41-20}$$

where

$$\alpha_n = j^n \sqrt{\nu_0 \lambda_n / x_0}, \tag{41-21}$$

and λ_n is a known constant (eigenvalue). Particularly, note in contrast to the direct Fourier transform that $G(\nu)$ is strictly bandlimited even though only a finite sample of $g(x)$ was used in determining a_n.

Clearly the PSWF expansions are very useful for determining the spectrum from a limited interferogram. In fact, if no noise is present, Eq. (41-18) provides all the resolution desired. On the other hand, if noise is present and the coefficient λ_n is carefully examined, it will be found that λ_n decreases as n increases. If any uncertainty (noise) exists in a_n, the higher order terms then become increasingly inaccurate. Thus, to prevent excessive errors in $G(\nu)$, the series must be truncated. Resolution is thus, not only limited, but is a function of the resulting spectrum SNR′. Rushforth and Harris[18] have examined the problem of various types of noise energy growth as n increases for the optical case. An exactly analogous set of calculations can be made for the four sources of noise in Fourier spectroscopy that have been mentioned. The results are the same as those of Rushforth and Harris:[18]

(1) Noise that is both bandlimited and truncated grows proportional to N, the largest value of n.

Background and detector noise are usually in this category.

(2) Noise that is truncated but not bandlimited, such as measurement noise, grows as

$$\sum_{n=0}^{N} 1/\lambda_n.$$

This is a much worse situation, as λ_n is very small for large n.

(3) Round-off error produces the most troublesome noise, which in general is neither truncated nor bandlimited. It grows as

$$\sum_{n=0}^{N} 1/\lambda_n^2. \qquad (41\text{-}22)$$

The increase in signal energy as n is increased depends, of course, on the exact nature of the incoming radiation, but since it is both bandlimited and truncated and must contain some structure, it can be expected to increase less rapidly than the bandlimited and truncated noise, thus satisfying the requirement that SNR′ must decrease as resolution increases. The importance of both filtering and truncating all noise sources whenever possible is well illustrated by these results.

It should be pointed out that for small retardation bandwidth cases (small c), the major increase in resolution gained using the PSWF technique is due to the conservation of information, not the sacrifice of SNR′ for resolution, although these two effects are mixed.

Figure 41-5 illustrates the utility of the PSWF technique. The two spectral bands are not observable when the finite FT is used but are readily apparent when the PSWF technique is applied. Unfortunately, such dramatic improvement cannot be obtained for large values of c. This is illustrated in Figure 41-6

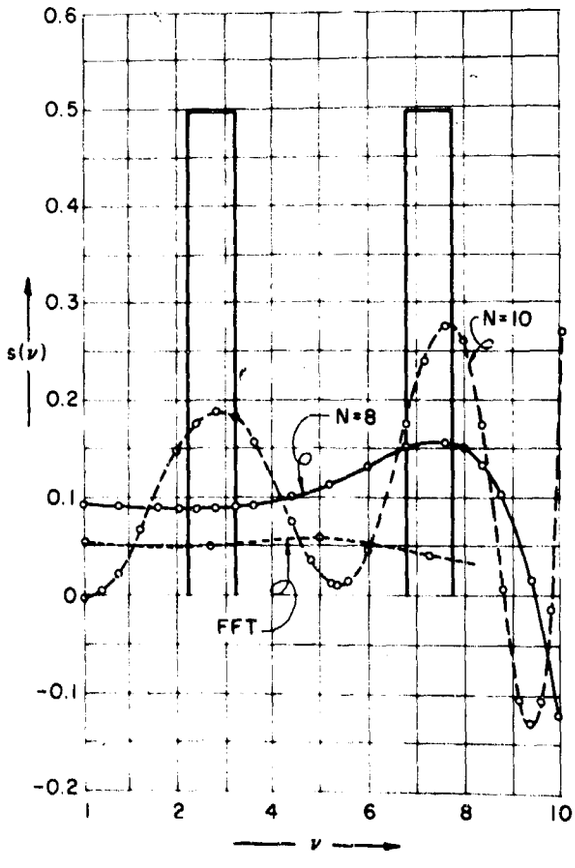


Figure 41-5. Input Spectrum and Spectral Estimates Using Finite FT and PSWF ($c = 10$)

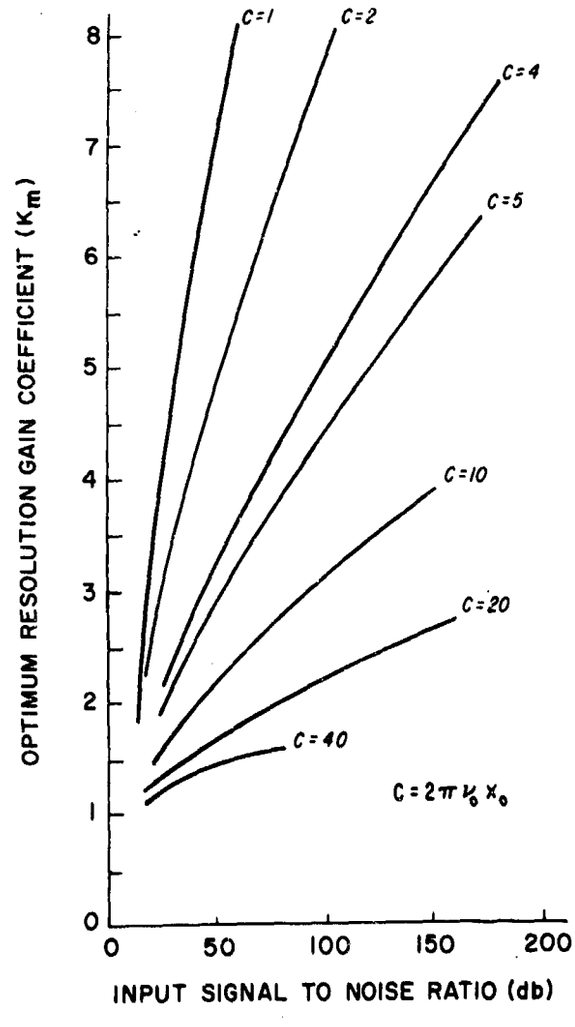


Figure 41-6. Optimum Resolution Gain, Minimum ISE Criterion, Narrow Spectral Band

where the optimum resolution gain is shown according to the integral squared error criterion.[17] Roundoff error becomes especially critical for large values of c, and computer calculations become increasingly difficult as c increases.

41-3.3 Method of Moments

The method of moments[20,21] is a technique of generalized expansion whereby the spectrum can be subdivided into separate pieces for processing. This has several advantages. First, *a priori* information can often be handled in a convenient way. Second, the effective retardation-bandwidth product is lower for a subsection than for the complete spectrum.

The problem to be solved is formulated in the following way:

$$F^{-1}[S(\nu)]=\hat{s}(x), \tag{41-23}$$

where $S(\nu)$ is to be found, $F^{-1}[\ ]$ is the inverse Fourier transform, and

$$\hat{s}(x)=\text{rect}\left(\frac{x}{2x_0}\right)[s(x)+n(x)]. \tag{41-24}$$

From Eq. (41-4) it is seen that $s(x)=d(x)-d(0)/2$.

First, the spectrum is divided symmetrically about $\nu=0$ into an even number of nonoverlaying subsections,

$$S(\nu)=\sum_{m=1}^{M} S_m(\nu-\nu_m), \tag{41-25}$$

each of a given bandwidth. Each of the subsections is then expanded in terms of the transforms of the PSWF with bandwidth corresponding to that of the subsection,

$$S_m(\nu-\nu_m)=\sum_{n=0}^{N_m} a_{nm}\Phi_{nm}(\nu-\nu_m), \tag{41-26}$$

where $\Phi_{nm}(\nu)$ is the Fourier transform of $\psi_n(x)$ with the bandwidth of the mth subsection. Then

$$S(\nu)=\sum_{m=1}^{M}\sum_{n=0}^{N_m} a_{nm}\Phi_{nm}(\nu-\nu_m). \tag{41-27}$$

Substituting this into Eq. (41-23) yields

$$\sum_{m=1}^{M}\sum_{n=0}^{N_m} a_{nm}F^{-1}[\Phi_{nm}(\nu-\nu_m)]=\hat{s}(x). \tag{41-28}$$

According to Fourier theory, this reduces to

$$2\sum_{m=1}^{L}\sum_{n=0}^{N_m} a_{nm}\psi_{nm}(x)\cos 2\pi\nu_m x=\hat{s}(x), \tag{41-29}$$

where $L=M/2$. Following the method of moments, the inner product of $f(x)$ and $g(x)$ is defined to be $\langle f(x), g(x)\rangle=\int_{-x_0}^{x_0} f(x)g(x)\,dx$; and the chosen weighting functions are $\psi_{jl}(x)\cos 2\pi\nu_l x$. Multiplying both sides of Eq. (41-29) by the weighting functions and forming the inner product, one finds

$$2\sum_{m=1}^{L}\sum_{n=0}^{N_m} a_{nm}\langle\psi_{jl}(x)\cos 2\pi\nu_l x, \psi_{nm}(x)\cos 2\pi\nu_m x\rangle = \langle\hat{s}(x), \psi_{jl}(x)\cos 2\pi\nu_l x\rangle. \tag{41-30}$$

This is more easily handled in matrix notation as

$$[\Omega_{jl,nm}][a_{nm}]=[s_{jl}], \tag{41-31}$$

where $\Omega_{jl,nm}=2\langle\psi_{jl}(x)\cos 2\pi\nu_l x,\ \psi_{nm}(x)\cos 2\pi\nu_m x\rangle$; and $s_{jl}=\langle\hat{s}(x),\ \psi_{jl}(x)\cos 2\pi\nu_l x\rangle$. The solution to Eq. (41-31) is

$$[a_{nm}]=[\Omega_{jl,nm}]^{-1}[s_{jl}]. \tag{41-32}$$

Using Eq. (41-27), one can then reconstruct an estimate of the spectrum. Notice also that for $M=1$ one has $\nu_m=0$, and this reduces to the simple case discussed previously of expanding the spectrum as a whole in terms of PSWF.

After straightforward, but lengthy, calculation,[22] one finds that for white noise of power spectral

density $N_0/2$, the following signal-to-noise energy ratio results:

$$\mathrm{SNR} = \frac{\sum_{m=1}^{L} \sum_{n=0}^{N_m} |a_{nm}|^2/\lambda_{nm}}{\frac{N_0}{2} \sum_{m=1}^{L} \sum_{n=1}^{L} \sum_{b=1}^{L} \sum_{i=0}^{N_m} \sum_{j=0}^{N_n} \sum_{a=0}^{N_b} \beta_{im,jn}\beta^*_{im,ab}\Omega_{jn,ab}/\lambda_{im}\lambda_{jn}} \tag{41-33}$$

where $[\beta_{im,jn}] = [\Omega_{jn,im}]^{-1}$, $\beta^*_{im,ab}$ is the complex conjugate, and λ_{nm} is λ_n for the value of c corresponding to the bandwidth of the mth subsection.

Thus, if a better approximation of the spectrum is found for $M > 1$ than for $M = 1$ where the final SNR′ is not less than the minimum tolerable SNR_m, there is effectively better resolution obtained.

Two examples are now given. In each the spectrum is approximated using $M = 1$ and $M = 2$. The spectra chosen are symmetric about $\nu = 0$, and only the positive parts are plotted in the accompanying figures. The power spectral density of the added noise is $N_0 = 10^{-7}$ W/Hz, and the smallest tolerable SNR was assumed to be $\mathrm{SNR}_m = 10^2$. Since the SNR is reduced in discrete jumps as more terms are taken, the final SNR′ in each case is the smallest possible such that it is greater than SNR_m.

In the first example, illustrated in Figure 41-7, $\mathrm{SNR}' = 3.6 \times 10^2$ for $M = 1$ and $\mathrm{SNR}' = 2.1 \times 10^2$ for $M = 2$. As seen in the figure, the better approximation results for $M = 2$. Figure 41-8 illustrates the second example, where $\mathrm{SNR}' = 1.3 \times 10^3$ for $M = 1$ and $\mathrm{SNR}' = 7.2 \times 10^4$ for $M = 2$. Here $M = 1$ yields the better approximation.

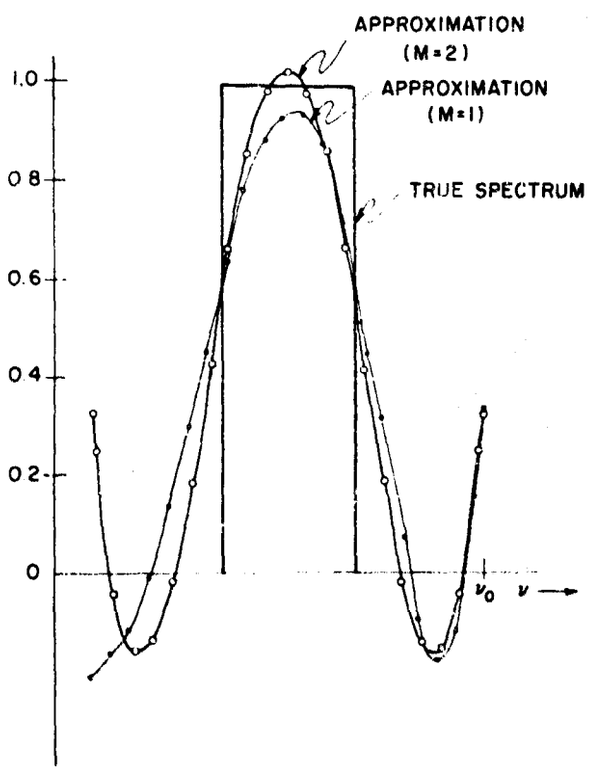


Figure 41-7. Method of Moments Technique of Resolution Enhancement ($m = 1, 2$; $c = 6$)

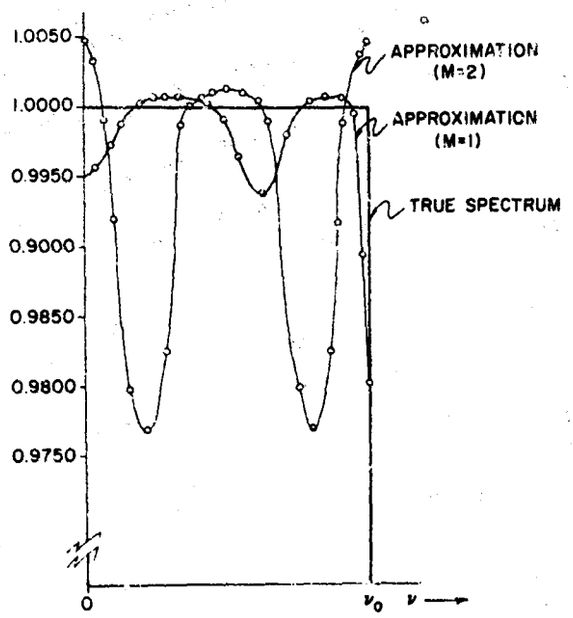


Figure 41-8. Method of Moments Technique of Resolution Enhancement ($m = 1, 2$; $c = 6$)

41-4 CONCLUSIONS

A theoretical relationship has been derived between the possible resolution enhancement and the resulting SNR that illustrates the large loss in SNR for significant increases in resolution (Figure 41-3). For a fixed length interferogram spectral resolution can be improved over that given by the finite FT, by using the PSWF. This improvement decreases as the retardation-bandwidth product increases. As more terms are added to the PSWF approximation, the noise increases as the resolution improves. If some *a priori* information is available about the spectrum, it may be worthwhile to consider the additional effort involved in applying the method of moments. It is especially important, in applying resolution enhancement techniques, that all system noise be suppressed, bandlimited, and truncated wherever it is possible to do so.

References

1. Michelson, A. A. (1891) *Phil. Mag.* **31**: 256.

2. Rubens, H., and Wood, R. W. (1911) *Phil. Mag.* **21**: 247.

3. Fellgett, P. B. (1951) Thesis, University of Cambridge.

4. Connes, J. (1963) NAVWEPS Report 8099, AD 409-869, U.S. Naval Ordnance Test Station, China Lake, Calif.

5. Loewenstein, E. V. (1966) *Appl. Opt.* **5**: 845.

6. Vanasse, G. A., and Sakai, H. (1967) *Progress in Optics*, Vol. 6, E. Wolf, Ed., North-Holland Publishing Co., Amsterdam.

7. Born, M., and Wolf, E. (1964) *Principles of Optics*, ed. 2 rev., Macmillan.

8. Shannon, C. E. (1949) *Proc. IRE* **37**: 10.

9. Shannon, C. E., and Weaver, W. (1949) *The Mathematical Theory of Communication*, University of Illinois Press, Urbana.

10. Woodward, J. C. (1969) Technical Report No. 6560-18, SV-SEL-69-039, Center for Systems Research, Stanford University.

11. Harris, J. L. (1964) *J. Opt. Soc. Am.* **54**: 931.

12. Forman, M. L. (1965) Private communication.

13. Barnes, C. W. (1966) *J. Opt. Soc. Am.* **56**: 575.

14. Flammer, C. (1957) *Spheroidal Wave Functions*, Stanford University Press, Stanford, Calif.

15. Williams, R. A., and Cheng, W. S. C. (1966) *J. Opt. Soc. Am.* **56**: 167.

16. Frieden, B. R. (1967) *J. Opt. Soc. Am.* **57**: 1013.

17. Kirlin, R. L., and Despain, A. M. (1968) Electro-Dynamics Laboratories Report 69-0034, Utah State University.

18. Rushforth, C. K., and Harris, R. W. (1968) *J. Opt. Soc. Am.* **58**: 539.

19. Slepian, D., and Pollak, H. O. (1961) *Bell Sys. Tech. J.* **40**: 43.

20. Harrington, R. F. (1967) *Proc. IEEE* **55**: 136.

21. Harrington, R. F. (1968) *Field Computation by Moment Method*, The Macmillan Company, New York.

22. Bell, J. W. (1970) Ph.D. dissertation, Utah State University.

Bibliography

Forman, M. L., Steel, W. H., and Vanasse, G. A. (1966) *J. Opt. Soc. Am.* **56**: 59.

Harris, J. L. (1966) *J. Opt. Soc. Am.* **56**: 569.

Harris, R. W., and Rushforth, C. K. (1966) Electro-Dynamics Laboratories Report 67-0070, Utah State University.

Helstrom, C. W. (1967) *J. Opt. Soc. Am.* **57**: 297.

Landau, H. J., and Pollak, H. O. (1961) *Bell Sys. Tech. J.* **40**: 65.

Landau, H. J., and Pollak, H. O. (1962) *Bell Sys. Tech. J.* **41**: 1295.

Wozencroft, J., and Jacobs, I. (1965) *Principles of Communication Engineering*, John Wiley and Sons, New York.

Discussion

Comment: (T. P. Sheahen, Bell Telephone Laboratories). There is a subtlety in the information-theoretic statement that "the computer only processes information, and doesn't generate information." While this is rigorously true, the computer is most often used to incorporate simple *a priori* knowledge in a very complicated way. This effectively enhances knowledge; the extensive information hidden away in the *a priori* knowledge is brought out by the computer.

Comment: (Dr. Pierre Connes, C.N.R.S.). It is highly interesting to see clearly for the first time how much we pay in signal-to-noise ratio for resolution enhancement; however, precisely because the price is so dear these schemes should be used only when resolving power is actually limited by the size or accuracy of optical elements. This includes spectra recorded with prisms, gratings or Fabry-Perot etalons. In this last case the plate separator is in practice limited by the finesse. The Fourier spectroscopy case is totally different; our high resolution interferometers (2 m path difference) use only techniques which could easily be extrapolated to any path difference. Thus the concept of instrument-limited resolving power is obsolete; the only limitation left is the true one—energy-limited resolving power. When planning today a Fourier spectrometer there is simply no excuse for not working the slides long enough. The use of railroad tracks, which is cheap and easily available, is respectfully suggested.

Comments: (Dr. A. T. Stair, Jr., AFCRL). Dr. Connes' comments do not apply to cases where the source is one-of-a-kind, time-limited, energy-limited, and particularly if it is *a priori* unpredictable. An example is a solar eclipse where a flare may occur in the space of a few seconds and measurements are being attempted for the first time at wavelengths and/or sensitivities heretofore not tried. Auroral phenomena and nuclear detonations are other examples where that particular event cannot be reproduced, but it is important, after the fact, to extract all of the information possible from the data. Furthermore, in many situations, e.g., satellites, size constraints are imposed.

Comments: (Dr. Lawrence Mertz, Smithsonian Astrophysical Observatory). I would like to call your attention to an article by Biraud (1969) *Astronomy and Astrophysics* 1: 124, which utilizes the valuable information that the spectrum is positive. Even if the augmentation of resolution should prove impractical, the academic understanding of mathematical processes is always worthwhile.

Contents

42. Feasibility of High Resolution Fourier Spectroscopy in the Vacuum Ultraviolet

A. S. Filler
623 Woodlawn Street
Fall River, Mass. 02720

Abstract

The design and operation of an interferometer for the region 500 to 1100 A is discussed. The instrument has the arrangement of a monochromator (such as Czerny-Turner, Littrow, or Pfund) with the dispersing element replaced by a pair of mirrors, one of which is moved to scan the interferogram. The use of multislits gives a large increase in throughput.

42–1 INTRODUCTION

Consider the problem of achieving high resolution (better than 0.1 cm^{-1}) in the vacuum ultraviolet, primarily between 500 and 1100 A (that is, between 0.9×10^5 cm^{-1} and 2×10^5 cm^{-1}). This region is called the helium ultraviolet (HeUV) since the high wavenumber limit is about where helium becomes opaque. The low wavenumber limit is chosen because then no materials are transparent in the region. The high wavenumber limit is chosen because it is the upper limit for normal-incidence reflection. These limitations restrict the design possibilities, and help one decide on the design.

Spectroscopists have long wanted a high resolution instrument in this region, and it was for this purpose that Michelson proposed the reflection echelon. But when it was finally determined how to construct this instrument, it proved so difficult to use that almost no useful work has been done with it. The difficulty comes from the very high order and restricted free spectral range. With 7 mm plates, the free spectral range is only 0.7 cm^{-1}, and at 700 A the order is 2×10^5. For an echelon with 40 plates the resolution then is about 0.02 cm^{-1}. Recently there have been great improvements in the production of gratings, but at present there are no gratings with the desired resolution in the HeUV. Work is now in progress on the production of gratings by holographic techniques. It remains to be seen how these perform in the HeUV. Even when better gratings become available there will remain a need for interferometric methods, just as there is in the visible and infrared, to provide wavelength standards, greater resolution, better scanning functions, and greater throughput.

42–2 GEOMETRY

The interferometers used for high resolution at smaller wavenumbers are the Michelson and the Fabry-Perot. Both require transparent plates to achieve division of amplitude and are not possible in the HeUV. Any interferometer in this region must use division of wavefront. A similar situation exists in the far infrared, and Strong introduced the lamellar grating interferometer for use in that region. It seems

improbable that the facets of the lamellar grating can be made sufficiently flat for use in the visible, let alone the ultraviolet. A usable interferometer results if the lamellar grating is reduced to a single pair of facets, which can consist of conventional plane mirrors, obtainable flat to at least 1/100 fringe of green light, or 1/9 fringe at 500 Å. This form of interferometer is called the Boltzmann interferometer. For strong modulation of the interferogram the slits must be narrow. If f is the focal length of the collimator, and d is the combined width of the two plane mirrors (assumed to be rectangular in shape), and the largest wavenumber with strong modulation is σ, then the slit width should be

$$w = f/\sigma d. \tag{42-1}$$

It is possible to use multislits with this interferometer.[1] The area that the multislits may cover is the same as the entrance and exit stops of interferometers with division of amplitude, namely, an annulus centered on the optical axis, with inner and outer radii r_1 and r_2 satisfying the relation

$$r_2^2 - r_1^2 = 2f^2/R = r_0^2. \tag{42-2}$$

The resolving power $R = \sigma/\Delta\sigma$ is given, for a two-beam interferometer, by

$$R = 2\sigma P, \tag{42-3}$$

where P is the path difference between the two beams.[2] Multiplying Eq. (42-2) by π shows that the areas of the annuli are constant, $S_0 = \pi r_0^2$. In the Michelson interferometer, the entrance and exit stops, or their virtual images, can overlap, and the entire annulus can be used. In an instrument with division of wavefront there are no virtual images, and if the entrance and exit beams are to be distinct, the entrance and exit stops must be exclusive. At most, each can consist of half the annulus. The division can be any diameter.

If the smallest wavenumber with strong modulation is σ/a, then the multislits can have a period

$$D = w(a+1). \tag{42-4}$$

With high resolution in the ultraviolet and a reasonable number of resolved elements, a is close to unity. For example, if $\sigma = 10^5$ cm^{-1}, $\Delta\sigma = 0.1$ cm^{-1}, and if there are 10^4 resolution elements then the spectral range is 10^3 cm^{-1}, and a is 1.01. The multislit array can then consist of clear and opaque strips of equal widths w. In the HeUV, w is small. Here, and later, dimensions of the instrument are $d = 5$ cm and $f = 2$ m. Then for $\sigma = 10^5$ cm^{-1}, w is 4 μ. With $\Delta\sigma = .02$ cm^{-1}, $r_0 = 1.27$ mm. With a Littrow or Pfund collimating system one can have $r_1 = 0$, $r_2 = r_0$. The central slit then has a length about 600 times its width. If the collimating system is a Czerny-Turner with annular stops, then slits tangent to the inner circle of the annulus also have a length $2r_0$. It will be difficult enough to make such narrow slits in short lengths. It may be desirable to sacrifice throughput by avoiding such long slits. This can be done fairly efficiently by using arcs of the annulus which are normal to the division between the mirrors. This is shown by the areas S_1 and S_2 in Figure 42-1. For the parameters given just above, if r_1 is 35 mm, then the width of the annulus, $r_2 - r_1$, is 23 μ, and the slits are only six times the width in length. If each arc has a length r_1, and if D has the value of Eq. (42-4), with a close to unity, then the open area of the entrance or exit stop is $S_0/12$. This may be compared with the value $S_0/500$ for a single slit of width w and length $2r_0$.

The use of a narrow annular stop has other advantages. The path difference depends on r, varying by a half wavelength, $\lambda/2$, between r_1 and r_2, and so the effective path difference P, which is an average over the area S, will depend on the nonuniformity of illumination with r. It will be easier to maintain uniform illumination over the small r-range of an annular stop than over a circular stop. If the illumination is not uniform over each radial direction, and if accurate wavenumbers are desired, then the nonuniformity must be evaluated and used as a weighting factor in determining the effective path difference. Note that the interferometer must be used in tandem with a low order grating monochromator in order to isolate a narrow spectral band. The grating instrument requires a narrow slit, which is conjugate to the stop S. This area matching can be done efficiently with the short annular stop, but not with the central circular stop.

For the interferogram to have strong modulation, the optical system must be diffraction-limited. This is difficult in this short wavelength region, but it should not be impossible. Mirror surfaces can be made with the required tolerances. No numerical calculations of aberrations have been made, but general considerations are encouraging. In particular, consider the Ebert collimating system (which is a special case of Czerny-Turner collimator with a single mirror) shown in Figure 42-1. This system has the advantage of being simple to adjust, and of automatically giving unit magnification when focused. The relative aperture will be small, $f/40$ with the parameters used before. This tends to reduce aberrations. The design also helps. With the plane mirrors P_1 and P_2 in the same plane, the system is symmetrical with unit magnification, and has zero coma and distortion. If the collimator C is a paraboloid there is no spherical

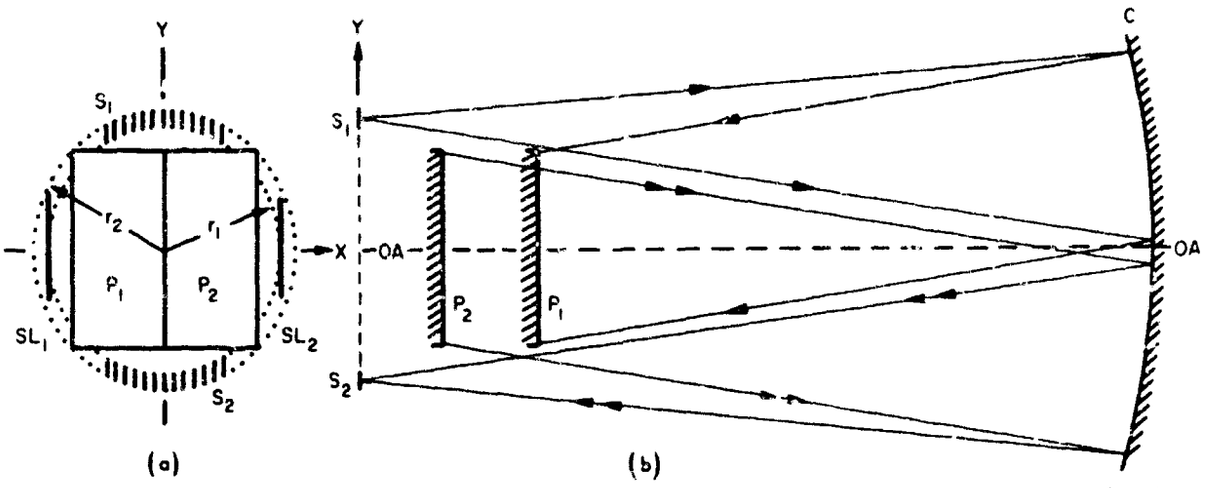


Figure 42-1. Schematic Diagram of a Boltzmann Interferometer with an Ebert Collimator (not to scale). (a) View from the center of the collimator C. (b) View normal to the Y-OA plane; P_1 and P_2 are plane mirrors; S_1 and S_2 are annular multislit arrays; Sl_1 and Sl_2 are alternative single slits; r_1 and r_2 are the radii of the annulus; C is a spherical or paraboloidal collimating mirror with optical axis OA, which focuses S_1 on S_2 at unit magnification. In order to avoid cluttering (b), only partial rays are shown: from S_1 to P_1 (single arrowhead) and from P_2 to S_2 (double arrowhead)

aberration, and if P_1 and P_2 are in the focal plane there is zero astigmatism. The only third-order aberration, then, is curvature of field, which is not disturbing. For reasonable dimensions the difference between a paraboloid and a sphere is small, on the order of 10 or 100 Å, and if it cannot be ignored, then the mirror figuring will be easy to achieve. In practice, one of the plane mirrors must be moved. It may not be convenient to have a mirror at the plane of the stops S, but especially for long focal lengths, these deviations from ideal configurations are pertubations, and the resulting aberrations will be small. Further investigation is required to find the minimum focal length for a given mirror size and positions for which the wave aberrations are less than $\lambda/4$.

42-3 MIRROR CONTROL

The position of the moving mirror must be well determined, and it must remain parallel to the fixed mirror. Tolerances are considered in the next section. At moderate wavelengths the effects of mirror tilts are avoided by using retroreflectors. These give two additional reflections, which are undesirable in the Hel'V. Interferometric control must be used for tilts as well as for position. A laser source is in many respects ideal for monitoring position, but commercially available lasers are not sufficiently stable. Their wavelength varies by about one part in 10^8, which is 10 Å in 10 cm. About one part in 10^{10} or 10^{11} is required. This can be attained by lasers stabilized against a molecular line.[2] In another paper,[3] a scheme is given for stabilizing any single-frequency laser against the center of a reference line with a stability of about one part in 10^{11}, depending on the power of the reference line. It employs a method of fringe setting which is also proposed for the mirror servo-controls. Instead of setting on the maximum or minimum of a fringe, two sets of fringes are produced in the servo-interferometer, with a relative shift of a half of one fringe width. The two sets of fringes thus intersect at regions of large slope, and the intersection of the shifted fringes is used for setting. Using reasonable parameters, namely a 6328 Å laser, 10^{-3} W, and a time constant of 10^{-4} sec, it is calculated that a precision in position of .01 Å can be had. Tilts can be controlled by using two additional position controls. A Michelson interferometer, both of whose mirrors are mounted on the moving mirror, can be used in zero-order with radiation from a tungsten lamp, and if the intersection of two shifted central fringes is used for setting, displacements due to a tilt can be kept within .02 Å, again using reasonable parameters. A more elaborate arrangement passes tungsten radiation successively through two Michelson interferometers, each of which has a mirror on one edge of the plane mirrors P_1 and P_2. This is similar to Ramsay's method[4] of controlling the parallelism of a Fabry-Perot interferometer, and has the advantage of automatically keeping P_1 and P_2 parallel to each other. It is not effective, however, at large path differences.[4]

The simplest way to install such interferometric control is to introduce collimated beams parallel to the direction of the mirror displacement. There may then be a systematic error in the position of the sampling points of the interferogram, since at different positions of the moving mirror its beam will fall on slightly different portions of the collimator. If the collimator had exactly the same path length for all rays this would do no harm. But the presence of aberrations, which can be calculated, or of figure errors, which must be measured, will introduce perturbations into the path difference. These must be evaluated, and if not negligible they must be corrected by interpolation of the measured interferogram.

42-4 SIGNAL AND NOISE

The casual reader may obtain the impression from the literature that Fourier spectroscopy offers no advantages when the predominant noise is due to photon statistics. It seems worthwhile to spend a little time examining this subject, if only to justify the rest of this paper. In general, note that photon statistics are always present, but are masked in the infrared by the less efficient detectors available there. It would be a paradox if a method were degraded just because the detector is improved.

Consider a spectrum consisting of a single narrow line, of wavenumber σ_1. The detector (photomultiplier) is considered as a photon counter, with a background counting rate B, which in practice should not be greater than about 1/sec. (The case of detector-limited noise is obtained by making B much larger than the signal counting rate.) Imperfections in the instrument will reduce the modulation of the interferogram, so that one must allow for a visibility V_1, which will be taken to be $\frac{1}{2}$ in the numerical examples. If R_1 is proportional to the peak counting rate, the interferogram after correction of phase errors, is

$$\begin{aligned} F(m\cdot\Delta L) &= (1+P_r)[Bt+R_1t(1+V_1\cos 4\pi\sigma_1 m\,\Delta L)] \\ &= (1+P_r)F_0(m\cdot\Delta L), \end{aligned} \tag{42-5}$$

where t is the time of measurement at each of the N sampled points with positions $L=m\cdot\Delta L$, and P_r is an operator which gives the noise in the interferogram, that is, gives a random quantity whose rms variation is $[F_0(m\cdot\Delta L)]^{\frac{1}{2}}$. The calculated spectrum is

$$S(\sigma_i)=F(0)+2\sum_{n=1}^{N-1}F(n\cdot\Delta L)\cos 4\pi\sigma_i n\,\Delta L \tag{42-6}$$

or

$$\begin{aligned} S(\sigma_i) &= F_0(0) \\ &\quad +2\sum_{n=1}^{N-1}F_0(n\cdot\Delta L)\cos 4\pi\sigma_i n\,\Delta L \\ &\quad +P_r[F_0(0)] \\ &\quad +\sum_{n=1}^{N-1}2\cos 4\pi\sigma_i n\,\Delta L P_r[F_0(n\cdot\Delta L)]. \end{aligned} \tag{42-7}$$

The first line in Eq. (42-7) gives the spectrum signal, a sine peak at $\sigma_i=\sigma_1$, and a spurious peak at $\sigma_i=0$ due to the terms $(B+R_1)t$ in Eq. (42-5). The spurious peak can be removed by subtracting from F_0 in the first line of Eq. (42-7) the term $(B+R_1)t$. The amplitude of the sine peak is $V_1NR_1t=V_1R_1T$, where $T=Nt$ is the total time of measurement. The total number of signal counts in the interferogram is R_1T, of which a fraction V_1 is effective in forming the spectrum.

The term $(B+R_1)t$ must not be subtracted from F_0 in the second line of Eq. (42-7) since this would remove most of the noise. The operator P_r produces random sign as well as random magnitude, and the random sign effectively removes the periodicity due to the cosine term in Eq. (42-5). Thus, the noise in Eq. (42-7) is independent of σ_i. If V_1 is close to unity the coefficient of R_1t in Eq. (42-5) can be approximated as a square wave, and the noise in Eq. (42-7) is the sum of about $N/2$ random quantities whose magnitudes have an rms variation of $[2\bar{c}(B+2R_1)t]^{\frac{1}{2}}$, where $\bar{c}=.64$ is the average of the absolute value of the cosine. The noise is $[\bar{c}(B+2R_1)T]^{\frac{1}{2}}$. If V_1 is small the cosine term in Eq. (42-5) can be neglected when calculating the noise, which then is the combination of N random quantities whose magnitudes have an rms variation of $[2\bar{c}(B+R_1)t]^{\frac{1}{2}}$. The noise is $[2\bar{c}(B+R_1)T]^{\frac{1}{2}}$. Except for very weak spectra, B is small compared with R_1 in the HeUV and both these noise levels equal $[1.3R_1T]^{\frac{1}{2}}$. At the spectrum peak the signal-to-noise ratio is about $0.9V_1[R_1T]^{\frac{1}{2}}$. The factor 0.9 can be set to unity without affecting the discussion.

If in addition to a strong single spectral element, as considered above, there is also a weaker line, also within a single spectral element, with a counting rate R', its presence will not change the noise level appreciably, and it will have a signal-to-noise ratio of about $V'R'[T/R_1]^{\frac{1}{2}}$. If, in order to be reasonably sure of not missing such a weak line because it is masked by the noise, or not mistaking a noise peak for a spectrum peak, the weak line has signal-to-noise ratio of at least three, then the useful range of intensity in the spectrum is $V_1(R_1T)^{\frac{1}{2}}/3$. If instead of a single strong spectral element there are M strong spectral elements, of average counting rate R_M, then the noise level in the spectrum is about $(MR_MT)^{\frac{1}{2}}$, and the useful intensity range is about $V(R_MT/M)^{\frac{1}{2}}/3$.

A criterion for the accuracy in positioning the moving mirror is found by allowing the error in the interferogram amplitude due to a position error e to be equal to (or less than) the uncertainty due to photon noise. Assume that

$$F_0(m\cdot\Delta L)=R_Mt\sum_{i=1}^{M}(1+V\cos 4\pi\sigma_i m\,\Delta L). \tag{42-8}$$

The worst error due to L being $m\cdot\Delta L-e$ will occur when all the cosine functions have the phase $\pi/2$

(or all $-\pi/2$). This extreme case is unlikely, but easy to calculate. Then the error in the interferogram is

$$R_M t \sum_{i=1}^{M} V 4\pi\sigma_i e = V 4\pi M R_M t \bar{\sigma} e,$$

where $\bar{\sigma}$ is the average wavenumber. The total photon count is $MR_M t$, and the photon noise is $(MR_M t)^{\frac{1}{2}}$. Therefore,

$$e = [4\pi\bar{\sigma}V(MR_M t)^{\frac{1}{2}}]^{-1} \qquad (42\text{-}9)$$

It may be useful to consider some numerical examples. Lacking data for ultraviolet sources, consider the calibration of the instrument by measuring the 0.6 μ line from a Kr-86 lamp. The slits, however, will be chosen to fit $\sigma = 10^5$ cm^{-1}, with $D = 4w$, so that $\frac{1}{4}$ of the entrance and exit stop areas are open. If the lamp bore has a diameter of 3 mm, and is imaged 1:1 on the entrance stop, the effective entrance area, or source area is .017 mm^2. The instrument accepts a solid angle $(5/200)^2$, or 1/1600 steradian. The source radiance[5] is 0.3 W · m^{-2} · sr^{-1}, so that $3 \cdot 10^{-12}$ W are available. Due to diffraction of .6 μ radiation at the 4 μ slits, only $\frac{1}{6}$ of the power incident on the entrance and the exit slits will be effective, reducing the power to about 10^{-13} W, or $3 \cdot 10^5$ photons/sec. With a trialkali cathode having a quantum efficiency of .05, and taking $V = \frac{1}{2}$, R is about 10^4 sec^{-1}. The line has a half-width[6] of .012 cm^{-1}, which is within the resolution limit of .02 cm^{-1} used in this paper. Taking $T = 10^3$ sec, the signal-to-noise ratio of the spectrum peak is 1500, and the intensity range is 500. If the interferogram is sampled at 10^3 points, t is 1 sec. The position tolerance e is 12 A. With the same parameters in Eq. (42-9), except with $\sigma = 10^5$ cm^{-1}, e is 1.6 A.

Now consider a continuous spectrum, again with $R_M = 10^4$ sec^{-1}, but with $M = 10^4$. With $T = 10^3$ sec, the signal-to-noise ratio is only 15, and the useful intensity range is 5. The value of e at $\sigma = 10^5$ cm^{-1} is .002 A, but since it is virtually impossible for all the phases in a continuous spectrum to coincide, a value of .02 A would be adequate.

42-5 OPERATION

The operation of a Fourier spectrometer in the HeUV differs from that in the infrared only in the nature of the detector and in the relative narrowness of the usable spectral band. Measurement of radiative power by photon counting has two advantages. A very low background can be obtained, and the measurements are directly in digital form. There is a practical limit to the counting rate, about 10^4 sec^{-1} to 10^7 sec^{-1} depending on the type of photomultiplier used. Near this upper limit a correction must be applied for pulse-coincidences, and above the limit the anode current must be measured, and put into digital form.

If the resolution is .02 cm^{-1} and a maximum of 10^4 spectral elements can be handled, the spectral range is 200 cm^{-1}. In the HeUV the only way to isolate such a narrow band is by using a grating monochromator. Unless suitable high-pass (in energy) filters can be found, perhaps in the form of thin foils, it will be necessary to operate the grating in the first order, so that lower energy radiation is not mixed with the desired spectral range. The higher energy radiation in higher orders should be easily removed using low-pass thin foil filters. A concave grating is preferred since it uses only one reflection. Such a grating, with 1800 lines/mm, and blazed at 3°, will be efficient near 600 A and at longer wavelengths, in the first order. If ρ, its radius of curvature, is 1 m the linear dispersion $d\sigma/dy$, at $\sigma = 10^5$ cm^{-1} is 550 cm^{-1}/mm, so that a band 200 cm^{-1} wide will be passed by an exit slit whose width is .36 mm. When used after the interferometer shown in Figure 42-1 the entrance slit of the monochromator is the exit stop S_2. The exit slit of the monochromator must be curved to match the curvature of S_2. If the latter has a radius and arc length both equal to 35 mm, then its sagitta is 5 mm. However, the grating also introduces some curvature into the image of its entrance slit, with a sagitta

$$s \simeq -\tan i(y^2/\rho), \qquad (42\text{-}10)$$

where i is the angle of diffraction, and x is the half-length of the slit. In the present example this is 0.3 mm, and the curvature of the exit slit must be changed accordingly.

The exit slit of the monochromator cannot be placed arbitrarily in its focal plane. Using the technique of 'aliasing,' a band $\delta\sigma$ wide need be sampled only at path difference intervals, ΔP, equal to $N/2\delta\sigma$, where N is the integer part of $\sigma/\delta\sigma$. In the example just considered, N is 500 and ΔP is 25 μ. But if this sampling interval is used, only spectra in intervals from $\sigma - \delta\sigma$ to σ will be aliased correctly. A band containing σ will show overlapping. Since ΔP is restricted to multiples of the wavelength λ used with the position-servo interferometer, there will always be wavenumbers which cannot be correctly aliased. The only solution is to have more than one source for the position interferometer. The avoidance of such 'forbidden wavenumbers' is made somewhat easier if ΔP is taken to be the wavelength of the interferometer source. There are two other advantages of doing this. An error in N of unity gives an error in σ

of $2\delta\sigma$. Using $\Delta P = \lambda$ gives N about 3 or 4, and an error in N is practically impossible. When N is 500 an error of one, although unlikely, is more possible. Recall that the analogous quantity for the echelon, the order, is $2 \cdot 10^5$, and then, apparently, an error of unity is difficult to avoid. When $\Delta P = \lambda$, a wide band spectrum can be calculated, at low resolution, and checked to see that there is no energy outside the desired band. The interferogram with the minimum number of points, according to the sampling theorem, can be obtained from the redundantly sampled interferogram by mathematical filtering.

When the monochromater is used after the interferometer, multichannel operation is possible by using more than one exit slit and detector. If many channels are used, this becomes expensive, but is many orders of magnitude less expensive than building as many interferometers.

42-6 CONCLUSION

The purpose of this paper is to decide whether or not an instrument with the proposed characteristics is feasible. There is no reason to believe that such an instrument will be built in the near future, and it seems premature to worry about the mechanical and electronic problems that will arise in engineering such an apparatus. The problems treated here are optical, and they appear to have satisfactory solutions.

A real instrument need not have the dimensions that were assumed, but they will not differ from these values by a factor more than 2 to 4. Fabricating the multislits may be difficult. Possible means include chemical etching of thin foils, using a photoresist which can be exposed through a lens or to the interference pattern of a laser, or by erosion using electron or laser beams. If satisfactory multislits cannot be made, then single slits ($S1_1$ and $S1_2$ in Figure 42-1) can be fabricated, but with a reduction in throughput of almost 1/100. This would make the instrument less versatile but it will still perform satisfactorily with sufficiently strong sources.

In the signal-to-noise ratio calculations, reflection losses at 0.6 μ were neglected; they will be relatively small. It is seen that to maintain a good signal-to-noise ratio with dense spectra, the integrated intensity RT of each spectral element must be high. This is due to the uniform distribution of noise in the spectrum, and not to the fact that photon noise predominated. Let us examine a case where detector noise predominates. This can be simulated by making the background noise in the interferogram, $(Bt)^{\frac{1}{2}}$, large compared with $R_M t$, the signal in the interferogram. Using $R_M = 10^4\ \text{sec}^{-1}$, which gave a signal-to-noise ratio of 15 in the earlier example with $M = 10^4$, and taking $t = 10^{-1}$ sec one needs $(Bt)^{\frac{1}{2}} \gg 10^3$, say $(Bt)^{\frac{1}{2}} = 10^5$. Then $B = 10^{11}\ \text{sec}^{-1}$. The signal-to-noise ratio is about $R_M(T/B)^{\frac{1}{2}}$ when V is one, and for $R_M = 10^4\ \text{sec}^{-1}$ the ratio is also one, so that the spectrum is useless. For the same conditions and a signal-to-noise ratio of 100, R_M must be increased to $10^6\ \text{sec}^{-1}$. Whether B is small or large compared with R_M, each spectral element (or, its sinusoidal modulation) is measured for the total time T, and both the advantages and disadvantages of multiplex operation are obtained.

It is interesting to compare the preceding examples with the performance of a monochromator with a hypothetical grating of equal resolution, or with a real echelon. For simplicity, assume that the grating has the same projected area as the split mirror. The angle i in Eq. (42-10) then has a tangent equal to five, and the sagitta equals the width w when the slit length is about 2 mm. Equation (42-10) is not a very good approximation at very large angles i, and the tolerable value of the sagitta will depend on whether or not the exit slit is curved and on how much i changes during a scan. Assume that the grating instrument has the same throughput as the interferometer with a single slit. Then R is about $2 \cdot 10^2\ \text{sec}^{-1}$. If there are 10^4 spectral elements to be scanned in 10^3 sec, the spectral signal is 20 counts, with a signal-to-noise ratio of five. The advantage of the grating instrument is its ability to measure weak lines in the presence of strong ones.

One tries to make the intensity at the detector as large as possible. But the signal-to-noise ratio depends on $VR^{\frac{1}{2}}$, for the case where photon noise predominates. If increasing R results in a decrease in V, and a decrease in $VR^{\frac{1}{2}}$, then the performance of the instrument will decrease. The use of rectangular mirrors gives the maximum area of the P mirrors with the maximum width of the slits. But it may be that rectangular mirrors cannot be made as flat as round mirrors, and this may decrease V, so the round mirrors may be preferable. The V term depends on the wavenumber, and the evaluation of $VR^{\frac{1}{2}}$ must be made in the region of interest.

The tolerance e on the mirror position, given by Eq. (42-9), depends on the time used to measure each point of the interferogram. Thus the tolerance can be increased by sampling the interferogram at more points. The value .01 A cited for the position error should be adequate for most cases. The tilt error is less critical, and a displacement due to tilt of .02 A is more than adequate.

Acknowledgments

The author would like to thank the Physics Department of the University of Utah, and the Alta Peruvian Lodge, for their hospitality during the preparation of this paper.

References

1. Filler, A. S. (1970) Multislits with Boltzmann and lamellar grating interferometers, to be submitted to *J. Opt. Soc. Am.*

2. Hall, J. and Barger, R. (1969) *Phys. Rev. Letters* **22**: 4.

3. Filler, A. S. (1970) Precision interferometric servo-control, submitted to *Appl. Opt.*

4. Ramsay, J. (1962) *Appl. Opt.* **1**: 411.

5. Hanes, G. R. (1963) *Appl. Opt.* **2**: 465.

6. Terrien, J. (1967) *J. Physique* **28**: C2, 3.

Contents

43. Extension of Fourier Spectroscopy to Centimeter Wavelengths*

J. Baukus and J. Ballantyne
School of Electrical Engineering
Cornell University
Ithaca, New York 14850

Abstract

A lamellar grating far-infrared interferometer has been used in the frequency range, 8 to 70 GHz. The instrumental modifications necessary to operate at this low frequency range are discussed, and measured spectra of some microwave sources are presented. The operation of a fast response time GaAs detector is reported.

43–1 INTRODUCTION

A lamellar grating far-infrared interferometer has been used to study the harmonic content of various microwave signal sources. Spectra of devices in the frequency range from 8 to 70 GHz (.27 to 2.3 cm^{-1}) were recorded. Some solid state microwave devices may have a high enough harmonic content to be suitable as sources for far-infrared spectroscopy. Also, knowledge of the harmonic content of these devices is useful in circuit design and in the study of electron transport phenomena. This paper will discuss the operation of the interferometer at these wavelengths and will present some preliminary results.

43–2 DESCRIPTION OF INTERFEROMETER

The interferometer, designed for use in the far-infrared spectral region,[1,2] contained a 35-cm square tempered aluminum lamellar grating with a grating constant of .43 cm, a 35 cm diameter, F:1.3 parabolic collimating and focusing mirror and lightpipe optics. The optical arrangement is sketched in Figure 43-1.

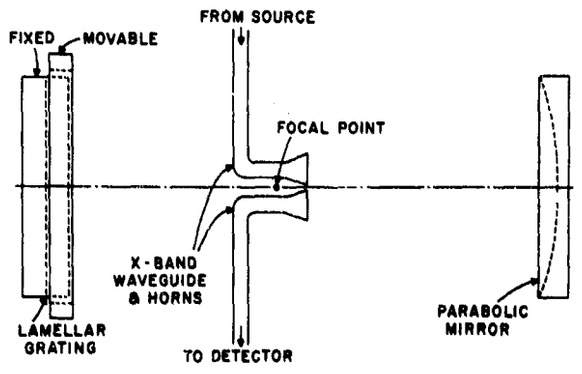


Figure 43-1. Diagram of the Optical Arrangement of the Lamellar Grating Far-infrared Interferometer

* This work was supported by the Air Force Systems Command, Rome Air Development Center, Griffiss Air Force Base, N.Y. with supplementary support from the Advanced Research Projects Agency through the Materials Science Center, Cornell University.

The grating, a portion of which is translated by a stepping motor and micrometer screw, splits the beam and introduces a path length difference. The beams are recombined at the exit slit producing a variation in output with path difference. This interferogram is the Fourier cosine transform of the incident radiation spectrum.[3]

Operation at microwave frequencies required a number of modifications to the interferometer. The 1.27 cm i.d. lightpipe was replaced by X-band waveguide as indicated in Figure 43-1. This permitted the use of X-band sources, because the cutoff frequency of the lowest order mode in the round lightpipe was about 15 GHz compared to 6.5 GHz for X-band waveguide. It was also found that the right angle bends in the lightpipe, which were formed with 45° first surface mirrors, caused severe reflections above the cutoff frequency. The X-band horns, used as the entrance and exit apertures, had the effect of lowering the maximum frequency to 500 GHz (about 17 cm^{-1}).[4]

Several other factors need to be considered in the operation of the instrument at these long wavelengths. The polarization of the radiation must be perpendicular to the lamellae of the grating as any parallel component will be reflected since the wavelength is comparable to the grating constant.[5] A piece of absorbing material was placed near the center of the parabola. Its size and position have a large effect on the minima of the interferogram. The resolution is limited by the thickness of the lamellae as the interference is destroyed at the longer wavelengths as soon as the movable portion of the grating becomes unmeshed from the fixed portion.

The detectors used were immersed in pumped liquid He at the end of a section of thin-walled stainless steel X-band waveguide. Detector elements were of two types, the first being a slice of a 2-W carbon composition resistor,[6] the second being a piece of epitaxial GaAs. They were mounted on microwave absorbing material at a nonperpendicular angle to the waveguide axis to prevent reflections from causing a standing wave or entering the interferometer. The GaAs detector was similar to that described by Stillman, et al.,[7] but was used in a new mode of operation at high bias which provided nanosecond response times for 10-GHz radiation. This detector is described in detail in another paper.[8]

43-3 SPECTRUM ERRORS

Operation of the interferometer with long-wavelength coherent sources posed a number of problems in obtaining good spectra. The principles of geometrical optics upon which the instrument was designed begin to break down due to diffraction effects at these wavelengths. Standing-wave and interference effects arise because of the coherence of the sources used.

Because the spectra obtained in this frequency range contained errors, the convolution method described by Forman et al.[9] was used to mathematically correct the interferograms. This method involves computing an error function from a short two-sided interferogram and convolving it with the data to obtain a corrected interferogram. The method produced only a slight improvement, probably because we could obtain only a few points behind the zero path difference position and thus did not get a good estimate of the phase error in the instrument. Some of the errors were also of a different nature which this method would not be expected to correct.

A periodic variation in transmitted power was observed when a large first surface mirror was taped to the grating and translated. Some of this variation was found to be due to stray reflections which were reduced by covering all exposed metal surfaces with microwave absorbing material (Emerson and Cumming AN73).

Since this did not eliminate the variation in transmitted power with the mirror placed over the grating, the possibility of Fresnel zone interference between the microwave horns was investigated.[10] This effect is due to interference produced at a point in space by the fixed phase relationship between points on a coherently illuminated aperture. Fresnel interference usually occurs at distances of a few wavelengths from the aperture, but it was felt that this effect might be significant in our case due to the imaging properties of the interferometer. In order to evaluate this effect, a computer program utilizing Kirchoff integrals to trace the radiation through the instrument was written. The program assumed a TE_{10} mode distribution of 1 cm wavelength radiating spherically from the entrance horn, and calculated the electric field at the exit horn aperture. The field distribution across a horizontal line at the center of the exit horn for various positions of an imaginary mirror in front of the grating is plotted in Figure 43-2. The distribution

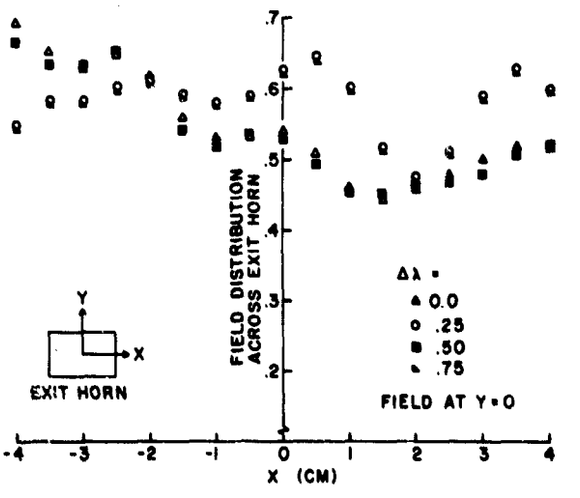


Figure 43-2. Computed Electric Field Distribution Across the Mouth of the Exit Horn for Several Values of Optical Path Length

varies with a period corresponding to a half wavelength change in the optical path between entrance and exit horns. The amplitude of the variation is comparable to the smallest variation in output power with mirror position that has been observed.

43-4 SPECTRA OF MICROWAVE SOURCES

The instrument performance, as indicated by spectra of various microwave sources, deteriorates with decreasing frequency. Figure 43-3 is the spectrum of a 65 GHz klystron. The indicated second harmonic power of about 4 percent of the fundamental was higher than expected, but this may be a harmonic or instrumental error. The 27 GHz klystron spectrum of Figure 43-4 indicates significantly more harmonic content. A section of V-band (50–75 GHz) waveguide, which acts as a 34 GHz high-pass filter was subsequently added, and no output was found. Thus, the indicated harmonic content is due solely to instrumental error. A later spectrum, taken after further addition of microwave absorbing material to the instrument, showed about half the harmonic content of Figure 43-4. The negative values near zero frequency are due to incorrect calculation of the average value of the interferogram.[11]

The spectrum of a 10 GHz klystron (Figure 43-5) indicates a large harmonic content which is also known not to exist. When a nonreflecting high-pass filter circuit was inserted between the klystron and the instrument, no output could be detected.

Inspection of the spectra of devices operating in the vicinity of 10 GHz indicated what appeared to be phase shifts in the interferograms. Because the convolution technique could not be used to correct these, the amplitude spectrum, that is, the square root of the sum of the squares of the cosine and sine transforms, is overplotted in Figure 43-5 for comparison. This type of spectrum appears, in some cases, to be a more reasonable approximation than the cosine transforms.

An X-band Limited Spacecharge Accumulation (LSA) device[12,13,14] spectrum (Figure 43-6) indicates a larger harmonic content than that of the klystron. The LSA device is a GaAs diode in which microwave oscillations are produced as a result of electron transfer between conduction band minima. When the high pass filter was inserted, power at the second harmonic was detected, and this filtered spectrum is overplotted in Figure 43-6. On an absolute scale, the peak value of the filtered spectrum corresponds to about 6 percent of the peak value of the full spectrum. The ratios of the second to third harmonics in the

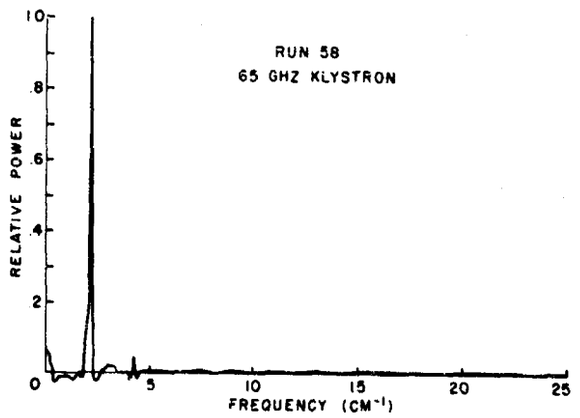


Figure 43-3. Spectrum of a 65 GHz Klystron (1 cm^{-1} = 30 GHz)

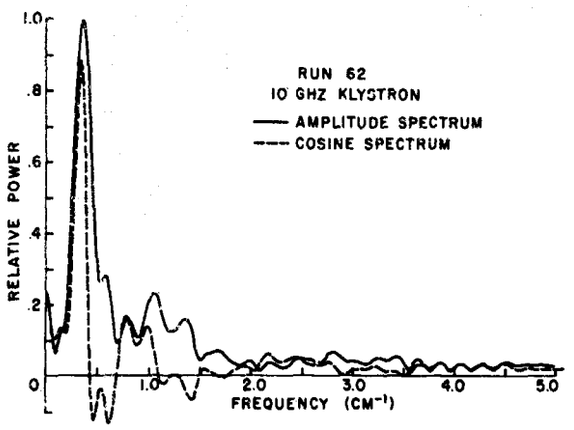


Figure 43-5. Comparison of the Amplitude and Cosine Spectra of a 10 GHz Klystron

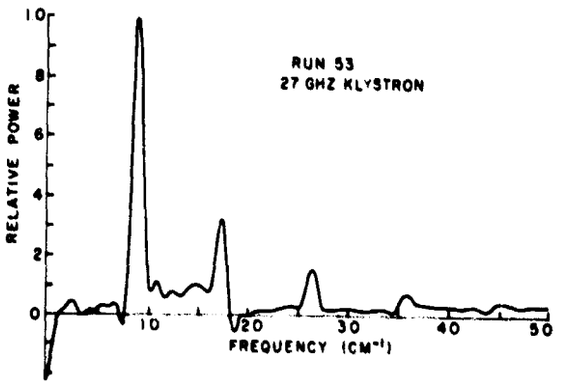


Figure 43-4. Spectrum of a 27 GHz Klystron

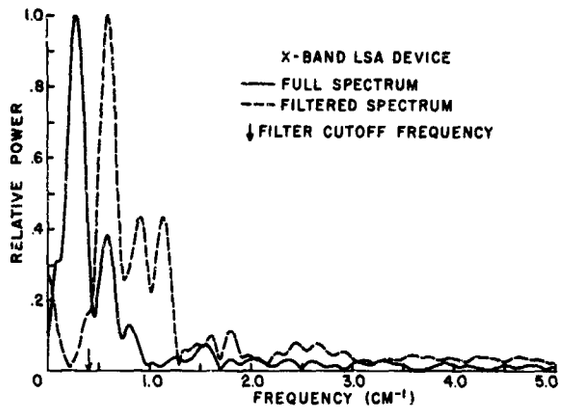


Figure 43-6. Spectra of an X-band LSA Device. The dashed line is the spectrum of the device operating through a nonreflecting high pass filter

filtered and unfiltered spectra are roughly equal. The large indicated fourth harmonic in the filtered spectrum is probably instrumental error as this is the second harmonic of the peak value.

The LSA devices were pulse operated to prevent destruction due to heating. The pulse widths were 1 μsec at 2000 pulses per sec. The data were recorded with the GaAs detector, several wide-band amplifiers, and a boxcar integrator.

43–5 CONCLUSIONS

While the use of this instrument at microwave frequencies is not without problems, it is a useful way to get information over a broad frequency range. The grating will successfully modulate radiation as low in frequency as X-band. Thus, when used with suitable high pass filters, information as to the magnitude and frequency of harmonics of microwave devices can be obtained, even without doing a Fourier transform. At higher millimeter wavelengths the instrument can be used, with some precautions, in its normal mode of operation. The use of the fast GaAs detector allows the investigation of pulsed devices. This investigation is not practical with the usual thermal bolometers.

Acknowledgments

The authors wish to thank Professor L. F. Eastman and Mr. J. Monroe for providing the LSA samples and circuits and for advice in operating them. The authors are grateful to the New York State Veterinary College, Physical Biology Computer Facility, which is supported by N.I.H. Grant No. FR-00326-2, for the use of its computer and computer programs.

References

1. Barnoski, M. K. (1968) Report No. 1003, Cornell Univ. Materials Science Center.

2. Barnoski, M. K. (1968) *Phys. Rev.* **174**: 946.

3. Strong, J., and Vanasse, G. A. (1959) *J. Opt. Soc. Am.* **49**: 844.

4. Richards, P. L. (1964) *J. Opt. Soc. Am.* **54**: 1474.

5. Gensel, L. (1965) *Japan. J. Appl. Phys.* **4**, Sup. 1: 353.

6. Boyle, W. S., and Rodgers, K. L., Jr. (1959) *J. Opt. Soc. Am.* **49**: 66.

7. Stillman, G. E., et al. (1968) *Appl. Phys. Letters* **13**: 83.

8. Baukus, J., and Ballantyne, J. (1970) to be published Proc. Symposium on Submillimeter Waves, Polytechnic Press, New York.

9. Forman, M. L., Steel, W. H., and Vanasse, G. A. (1966) *J. Opt. Soc. Am.* **56**: 59.

10. Silver, S. (1949) *Microwave Antenna Theory and Design*, McGraw-Hill, New York, p. 169.

11. Blackman, R. B., and Tukey, J. W. (1959) *The Measurement of Power Spectra*, Dover, New York, p. 95.

12. Kennedy, W. K., Jr., and Eastman, L. F. (1967) *Proc. IEEE* **55**: 434.

13. Copeland, J. A. (1966) *Proc. IEEE* **54**: 1479.

14. IEEE Trans. Elec. Dev. (1967) ED-14, Sept. 1967.

Contents

44. Signal-to-Noise Ratio in a Photographic Fourier Transform

P. Bouchareine and R. Chabbal
Laboratoire Aimé Cotton, Faculté des Sciences
91 Orsay, France

Abstract

The conventional spectrograph is compared with a photographic interferometer where fringes are recorded instead of dispersed spectral lines. It is shown that if preexposure is not required, a photographic Fourier transform produces a gain in exposure time but with a corresponding loss in signal-to-noise ratio in the same manner as a photomultiplier.

44-1 INTRODUCTION

Multiplex methods give an important improvement in the signal-to-noise ratio of a spectrum when the noise is produced only by the receiver. This advantage vanishes in some cases when the noise is caused by photon noise in the signal that impinges on the receiver. In the analysis by R. C. Jones,[1] the photographic receiver was considered to produce a noise with the same properties as infrared cells. Consequently, the photographic detector might be able to give a significant enhancement in signal-to-noise ratio if a multiplexing method is used.

If preexposure is not required, a photographic Fourier transform produces a gain in exposure time but with a corresponding loss in signal-to-noise ratio in the same manner as a photomultiplier.

44-2 SIGNAL-TO-NOISE RATIO IN A PHOTOGRAPHIC DENSITY MEASUREMENT

The response of a photographic emulsion is not linear, and the signal-to-noise ratio is not directly observable on a densitometric trace. Let γ be the slope of the Hurter-Driffield curve

$$D = \gamma(\log n), \tag{44-1}$$

where n is the number of photons received per unit area on the plate.

The uncertainty δn that corresponds to an error δD of the density measurement is (Figure 44-1)

$$\frac{\delta n}{n} = \frac{\delta D}{\gamma}. \tag{44-2}$$

In the following approach, a well-known result for the error δD_s in a measurement of an area S of the plate is used. This is given by Selwyn's law[2] which states that δD for a given density is proportional to

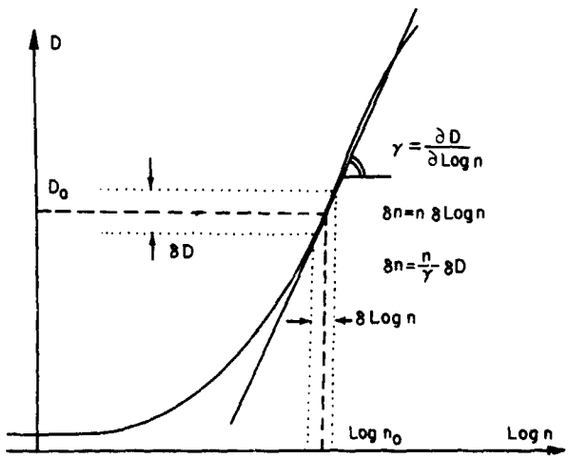


Figure 44-1. Derivation of Signal-to-Noise Ratio $n/\delta n$ from Hurter and Driffield's Curve

the inverse of the square root of S. If δD is the error for a unit area of the plate caused by the granularity, the error for an area S is then

$$\delta D_s = \frac{\delta D}{\sqrt{S}}. \tag{44-3}$$

This error is translated into an uncertainty in the measurement of the total number of photons received on the area S by

$$\delta N = S \times \delta n = \frac{S \times n}{\gamma} \frac{\delta D}{\sqrt{S}} = \frac{n}{\gamma} \delta D \sqrt{S}. \tag{44-4}$$

The signal-to-noise ratio will then be

$$\rho = \frac{N}{\delta N} = \frac{nS}{\delta N} = \frac{\gamma}{\delta D} \sqrt{S}. \tag{44-5}$$

The value of δD depends on the average density D. It is necessary, in a given experiment, to choose the exposure time or the illuminated area of the plate in such a way that the final density will be D_0, for which the efficiency (more exactly the detective quantum efficiency[3,4]) of the plate has its highest value. Experimental comparison only is possible between two experiments done with two plates of different characteristics, or with two identical plates used at different densities, but a theoretical comparison is possible between two experiments using the same emulsions and working at the same density. The signal-to-noise ratios are calculated from Eq. (44-5) and the exposure times are derived from photometric data, assuming validity of the reciprocity law.

44–3 COMPARISON BETWEEN A CONVENTIONAL SPECTROGRAPH AND A PHOTOGRAPHIC INTERFEROMETER

Let us consider a conventional spectrograph that uses a prism or a grating as the dispersive element. The exit pupil area is S_i. The output beam, divided into k equal parts corresponding to k spectral elements, is incident on a plate of area S. If F is the focal length of the system, then the output optical acceptance U is given by

$$U = \frac{S \cdot S_i}{F^2} = S \cdot \Omega, \tag{44-6}$$

where Ω is the solid angle of the beam impinging on the plate. Since the beam is dispersed into k spectral elements, the input optical acceptance is only (Figure 44-2)

$$U_i = \frac{U}{k}. \tag{44-7}$$

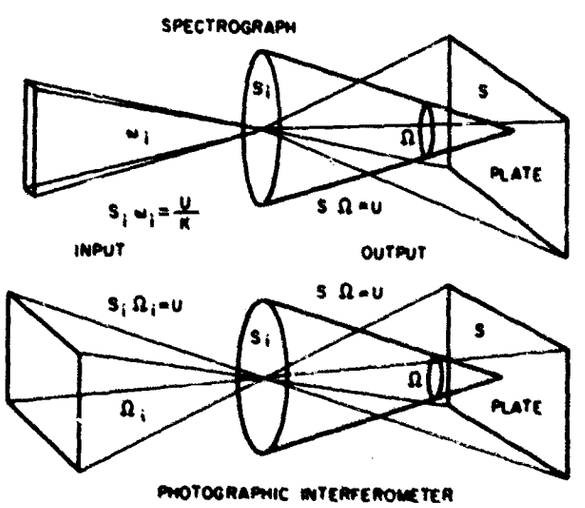


Figure 44-2. Comparison between Geometrical Parameters of the Interferometer and the Spectrograph

Let us now consider a two beam interferometer producing, in the same area S of an identical plate, straight-line fringes which are analyzed later by a Fourier transform method, as has been suggested by several authors.[5,6,7] In order to have instruments that can be compared, we take the same value of the plate area S, of the pupil area S_i and of the solid angle Ω of the output beam. In this case, the input optical acceptance for the interferometer is the same as the output acceptance U. Furthermore, it is not spatially divided by any disperser.

44-3.1 Single Line Case.

44-3.1.1 SPECTROGRAPH

Let B equal the number of photons emitted by the source per unit time and unit optical acceptance, and T equal the exposure time. Then the number of photons n received on the spectrograph plate per unit area of the slit image will be

$$n = B \cdot \Omega \cdot T. \tag{44-8}$$

If n_0 is the value of n that gives the optimum density D_0, the corresponding exposure time is

$$T_0 = \frac{n_0}{B \cdot \Omega}. \tag{44-9}$$

According to Eq. (44-5), the signal-to-noise ratio obtained by measuring the density at the slit image will then be

$$\rho_s = \frac{\gamma}{\delta D_0}\sqrt{\frac{S}{k}}. \tag{44-10}$$

In Eq. (44-10) k is the number of spectral elements that could be present, but there is only a single line in this case.

44-3.1.2 PHOTOGRAPHIC INTERFEROMETER

As can be seen from Eq. (44-8), the illumination of the plate in a bright fringe is the same as that of the slit image in the spectrograph*, and the time needed to obtain the density D_0 is T_0.

The density measurement of D_0 is done on an area $S/2$ of the plate.† The signal-to-noise ratio is therefore

$$\rho_I = \frac{\gamma}{\delta D_0}\sqrt{\frac{S}{2}}. \tag{44-11}$$

44-3.2 Continuous Spectrum

Let us assume now that the k spectral elements have equal luminance B. There is actually a spectral luminance $B(\lambda)$ which is a constant determined by the condition

$$B = \int_{\delta\lambda} B(\lambda)\, d\lambda. \tag{44-12}$$

44-3.2.1 SPECTROGRAPH

The number of photons received by the plate is the same as in the previous case. The exposure time is T_0 and since each spectral element is independent of the others, the signal-to-noise ratio has the same value:

$$\rho_s = \frac{\gamma}{\delta D_0}\sqrt{\frac{S}{k}}. \tag{44-13}$$

44-3.2.2 PHOTOGRAPHIC INTERFEROMETER

In the interferometer, the number of photons received per unit area of the plate is multiplied by $k/2$, because the k spectral elements overlap over the entire area of the plate, and they are transmitted with a factor $\frac{1}{2}$ that arises from the interference fringes. The exposure time needed to obtain the average density D_0, assuming the validity of the reciprocity law, is

$$T = \frac{2T_0}{k}. \tag{44-14}$$

The signal corresponding to each spectral element is now the amplitude of a density pattern of a given spatial frequency. Assuming once more a square modulation, this signal ΔD is the difference between two densities

$$D_1 = D_0 + \frac{\Delta D}{2} \text{ and } D_2 = D_0 - \frac{\Delta D}{2} \tag{44-15}$$

* This is true for a Michelson interferometer in which half of the light goes back to the source. In some types of interferometers where the brightness of the fringes is twice that of this one.

† This is true if a square modulation is assumed. For a sinusoidal pattern the signal-to-noise ratio is difficult to calculate, but the numerical results do not significantly differ.

each measured on a plate area $S/2$.

The density error for each one is

$$\delta D = \delta D_0 \sqrt{\frac{2}{S}}, \tag{44-16}$$

and the signal-to-noise ratio is

$$\rho_I = \frac{\Delta D}{\delta D} = \frac{\Delta D}{\delta D_0} \frac{\sqrt{S}}{2}. \tag{44-17}$$

The signal amplitude D is given by the modulation of a given spectral element. As n_0 is the average total number of photons received per unit area of the plate, this modulation is given by

$$\Delta n = \frac{2n_0}{k} \tag{44-18}$$

photons per unit area. Hence, the signal ΔD is

$$\Delta D = \gamma \frac{\Delta n}{n_0} = \frac{2\gamma}{k}, \tag{44-19}$$

and then the signal-to-noise ratio is

$$\rho_I = \frac{\gamma}{\delta D_0} \frac{\sqrt{S}}{k}. \tag{44-20}$$

A comparison of this last expression with Eq. (44-13) shows that, when the spectrum is continuous with constant spectral luminance $B(\lambda)$, the interferometric method (which provides a greater amount of light and consequently allows a shorter exposure time) leads to a signal-to-noise ratio for one spectral element roughly that of the conventional spectrographic method divided by the square root of the gain of exposure time. This result might be achieved in the same conventional spectrograph with the use of a camera lens of $\sqrt{k}$ times greater aperture, that is, with the same pupil area and a $\sqrt{k}$ times shorter focal length. The results of the comparison of the two instruments are given in Figure 44-3.

		INTERFEROMETER	SPECTROGRAPH
OPTICAL ACCEPTANCE		U	U/k
FLUX PER UNIT AREA OF THE PLATE	1	$B_\lambda \Omega$	$B_\lambda \Omega$
TIME EXPOSURE	1	$T = \frac{n_0}{B_\lambda \Omega}$	$T = \frac{n_0}{B_\lambda \Omega}$
SIGNAL TO NOISE	1	$(\gamma/\delta D)\sqrt{S/2}$	$(\gamma/\delta D)\sqrt{S/k}$
1 MONOCHROMATIC	2	$B_0 \Omega k/2$	$B_0 \Omega$
2 CONTINUOUS	2	$2T_0/k$	T_0
	2	$(\gamma/\delta D)\sqrt{S/k}$	$(\gamma/\delta D)\sqrt{S/k}$

Figure 44-3. The Results of Comparison Between the Two Instruments

44-4 CONCLUSIONS

Recording fringes instead of dispersing lines allows shorter exposure times for a given plate area and a given solid angle of the output beam, but this operation reduces the signal-to-noise ratio available from the plate if there are many spectral elements in the spectrum.

This result can be interpreted by the quantum nature of the photographic noise; the noise being produced by the photon noise of the whole spectrum while a single spectral element is measured. This conclusion agrees with a more general result obtained[9] as a direct consequence of Selwyn's[2] law. The noise observed in a photographic density measurement for the optimum value of the density is determined by the same laws as the photon noise observed with a photomultiplier. Findings of this investigation do not disagree with Jones's classification because he studied the detectivity of a plate, defined by the minimum detectable power for a given exposure time. Under such conditions it is necessary to preexpose the emulsion, and the noise produced by the parasitic beam used for the preexposure acts like the constant fluctuations of a noisy detector. The consequence for multiplex spectroscopy with photographic detection is quite similar to findings in an experimental study of photographic grid spectroscopy.[10]

However, such multiplex use of a plate has several advantages.

(1) If preexposure should be necessary, it is preferable to use a grid or a multiplex method because preexposure is then achieved with the light of the spectral elements being studied, instead of using a parasitic beam which simply contributes noise.

(2) Even if theoretical improvement in signal-to-noise ratio cannot be expected (when many elements are present), it is not possible with a great number of lines of different brightness to obtain in a single exposure the optimum value of density for each of them. This is possible with an interferogram for which (the zero path difference excepted) the illumination of the plate is much more homogeneous. Such a low dynamic range can be obtained with better linearity of the emulsion's response.

(3) A distinct advantage results when the spectrum has few spectral elements: the slit of a spectrograph uses only a small part of the area of the plate. In such a case, the interferometer makes better use of the information capacity of the plate in addition to achieving a gain in the transmitted flux.

This investigation was done as part of a contract with the "Direction des Recherches et Moyens d'Essais."

References

1. Jones, R. C. (1953) *Adv. in Electr. V*, Acorn Press, New York. pp. 1, 96.

2. Selwyn, E. W. H. (1935) *Phot. J.* **75**: 571.

3. Fellgett, P. B. (1958), *Month. Not. of Roy. Astr. Soc.* **224**: 3.

4. Jones, R. C. (1958) *Phot. Sc. Eng.* **2**: 57.

5. Lowenthal, S. Optics Communications, to be published.

6. Prat, R. *Opt. Act.*, to be published.

7. Stroke, G. W. (1967) *J. Phys.* **28**, C2.

8. Yoshihara, K., and Kitade, A. (1967) *Japan. J. Appl. Phys.* **6**: 116.

9. Bouchareine, P. *Opt. Act.*, to be published.

10. Bouchareine P., and Jacquinot, P. (1967) *J. Phys.* **28**, C2.

Contents

45. Devices for Multiplex Stellar Spectroscopy

André Girard
Office National d'Etudes et de Recherches Aérospatiales (ONERA)
92-Châtillon, France

Abstract

Two devices for stellar spectroscopy are described: a conventional disperser type (with grille) and an interferometer with no internal moving parts. Both of these devices have potentials for space applications. Preliminary results of blackbody radiation obtained with the grille spectrometer are presented. The first stellar spectrum of αORI obtained with the interferometer from the Jungfraujoch observatory is also presented.

45-1 INTRODUCTION

Of the two well-known advantages of Fourier spectroscopy (the spatial one of high optical acceptance, and the time one of multiplex operation), the second one is of particular importance in the case of stellar spectroscopy, as the optical extent (or *étendue*) of the beam is generally much smaller than the optical acceptance of an ordinary interferometer. Simple devices can thus be designed for stellar spectroscopy, especially if the requisite number of spectral elements is not very large.

Even including turbulence phenomena, the throughput (Jacquinot) advantage or space advantage given by an interferometer is not generally fullfilled in stellar spectroscopy. The other one, the multiplex (Fellgett) advantage or time advantage is intrinsic in this case as usual.

Both devices described here have been studied with this in mind and also with a view toward applications where the required number of spectral elements is moderate (a few hundred). Under these conditions, the expected value of multiplex advantage is between 10 and 30. Thus, the price paid for this gain must not be too high. Simple devices are required.

45-2 MULTIPLEX GRILLE SPECTROMETER

The multiplex grille spectrometer uses the optical working principle of a standard spectrometer. Upon dispersion by a prism or a grating, the star spectrum is linearly displayed in the focal plane of the exit collimator. It is now intended to investigate the spectral intensity distribution along this line. The problem considered can be treated as a special case of a multiplex image analysis method.[1,2,3]

The linear distribution of luminance is analyzed by modulating each point of the line with a time frequency, itself a linear function of the point abscissa on the line. This is achieved by placing, in the same plane, a grille with alternate opaque and transparent zones limited by equilateral hyperbolas (Figure 45-1). The grille is translated at uniform speed perpendicularly to the line under analysis. This linear motion can be converted into a rotating motion by using an appropriate zone distribution.[2,3]. The signal, which is a function of time, is the sum of elementary sinusoidal contributions from all spectrum elements along the line.

The principle underlying this spectroscopic device is similar to the theory governing interferometric

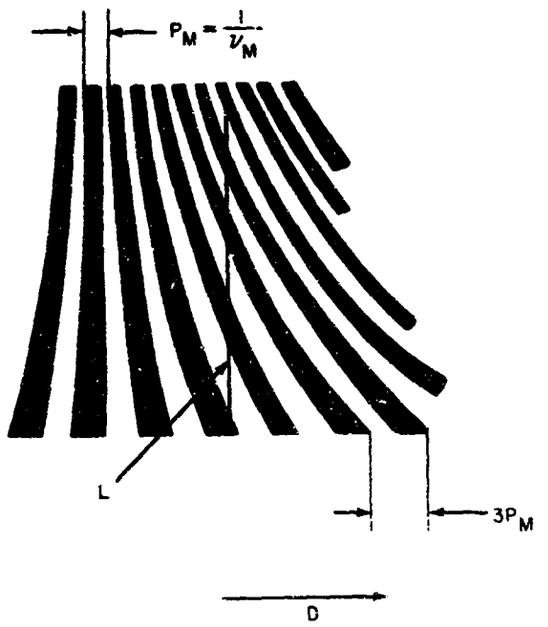


Figure 45-1. Grille Showing: D=Scanning Direction, L=Line Under Analysis

Fourier spectroscopy. For instance, the zero-path-difference position is obtained when the common asymptote of the set of hyperbolas is superimposed on the line under study. However, some specific problems do arise here. In practice, the transparency law governing the grille is not a sinusoidal law, so that odd harmonics of the main frequency occur in the signal. This difficulty is overcome by limiting the fundamental frequency range to the interval lying between ν_M, the maximum frequency, and $\nu_M/3$ (Figure 45-1). Thus, the first harmonic of the lower main frequency (that is, $3-(\nu_M/3)$), lies just outside the useful frequency range. It is easy to demonstrate that the number of discrete spectral elements is

$$N = \frac{4}{3}\frac{\Delta_M}{P_M}, \qquad (45\text{-}1)$$

where Δ_M is the maximum value of translation and P_M is the smallest interval of the grille.

The total resolving power of the system depends upon (and is at most, equal to) the resolving power of the spectrometer preceding the grille; it can be very high for a narrow spectral interval.

The optical étendue of the beam emerging from the telescope is greatly enlarged by the dispersing system. This can be a major disadvantage when a high-sensitivity, cooled infrared detector is used. This drawback is obviated by a second subtractive dispersion equal to the first; in this way the optical acceptance of the beam is restored to its initial value.

The main advantage of this device lies in the fact that the margin of tolerance required for adjustments and mechanical motions lies within a fraction of the smallest step P_M of the grille instead of covering a part of the wavelength as in the standard interferometric case.

The main characteristics of the experimental device under review are as follows:

Spectral interval: 1 μ to 2 μ
Disperser: Flint prism
PbS cell: 1 mm^2
$\Delta_M = 30$ mm
$P_M = 0.2$ mm
$N = 200$

The apodized instrumental function is shown in Figure 45-2. Residual defects are probably due to irregularities in the original drawing of the grille.

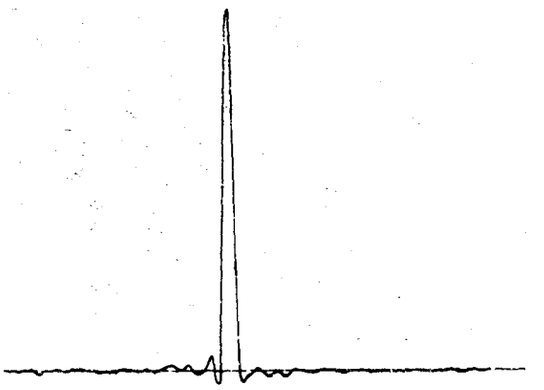


Figure 45-2. Apodized Instrumental Function

Figure 45-3(a) shows the record and Figure 45-3(c) shows the computed spectrum for a blackbody source at 900°C. Between these two curves Figure 45-3(b) shows the spectrum obtained by the conventional slit method, in the same length of time and with equal resolving power.

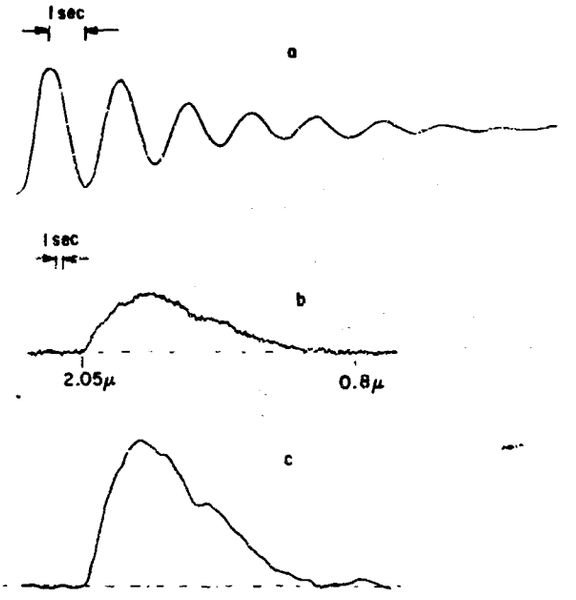


Figure 45-3. Blackbody Source at 900° C; a: multiplex record, b: recorded spectrum obtained with a slit, c: computed spectrum from curve 'a'

45-3 "AIR WEDGE FRINGES" INTERFEROMETER

In stellar spectroscopy, the optical acceptance of an interferometer is generally much larger than is necessary for the beam from the telescope. For instance, the optical acceptance of a Michelson interferometer equipped with mirrors 50 mm in diameter is 7 mm rad for a path difference of 100λ, as compared with 0.09 mm rad for the beam from a 600 mm diameter telescope of low optical quality (30 sec of arc).

Accordingly, a standard interferometer is too powerful an instrument for stellar spectroscopy. This applies even more to field widening facilities.[4] Thus, in a stellar interferometer, the useful area of the mirrors can be reduced to a very small portion of their actual area. This characteristic is used in the interferometer described below.

In this unit, which is of the Michelson type, the plane mirror in one arm is of a fixed off-axis design, consistent with the well-known "air wedge fringes" adjustment.[5] Thus, the path difference changes linearly in the plane of localized fringes. With the star image focused in this plane on an area smaller than a half-interfringe, the Fourier transform of the spectrum is obtained by sliding the interferometer as a whole perpendicularly to the system of equidistant fringes.

The interferometer incorporates no moving part; it is a compact solid rectangular block of four bonded elements (Figure 45-4). Internal elements 1 and 2, arranged on either side of the beam-splitter coating S, are made of the same material (refractive index n), while external elements 3 and 4 are made of another material (refractive index n'). Figures 45-5(a) and 45-5(b) show the end cross sections of the interferometer. As can be seen, each of elements 2 and 4 has equal respective cross sections all along its length, while elements 1 and 3 vary in thickness. The values of e and l, thickness differences in the two end cross sections respectively for elements 1 and 3, are defined by two elementary conditions:

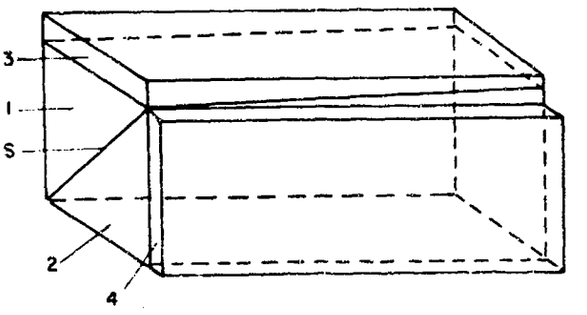


Figure 45-4. "Air Wedge Fringes" Interferometer

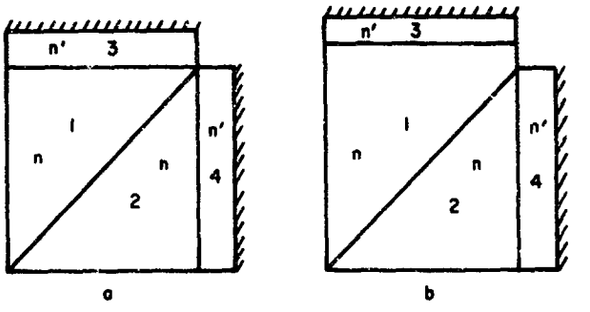


Figure 45-5. "Air Wedge Fringes" Interferometer. End cross sections

$$nl+n'e=\Delta/2$$

and

$$n'l+ne=0, \qquad (45\text{-}2)$$

where Δ is the path difference required. The second equation is the field widening condition. In this way, a very large aperture is obtainable for any path difference.

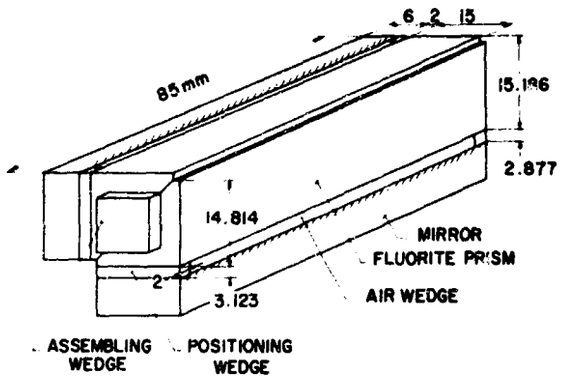


Figure 45-6. Fluorite-air "Air Wedge Fringes" Interferometer

In the interferometer used for stellar observations (Figure 45-6), field-widening compensation is achieved by combining a fluorite prism and an air wedge. The measurements of the interferometer are 18×18×80 mm. The resolution obtainable is 300 at 2 μ.

The stellar spectrometric device as a whole will be described in a later publication. Figure 45-7 shows the apodized instrumental function. Figure 45-8 represents the first stellar result obtained. This spectrum of αORI was produced at the Jungfraujoch observatory in Switzerland using a 760 mm diameter telescope. It is uncorrected for sky emission or atmospheric transmission.

The main feature of this interferometer is its foolproof safety against misadjustment. While this is no major advantage for observatory work, it is of interest in space experimentation, which has been the object in view in this particular case.

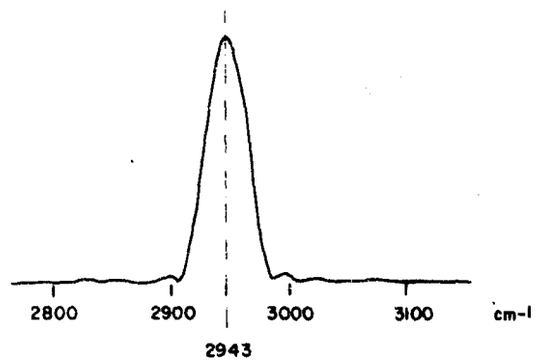


Figure 45-7. Apodized Instrumental Function

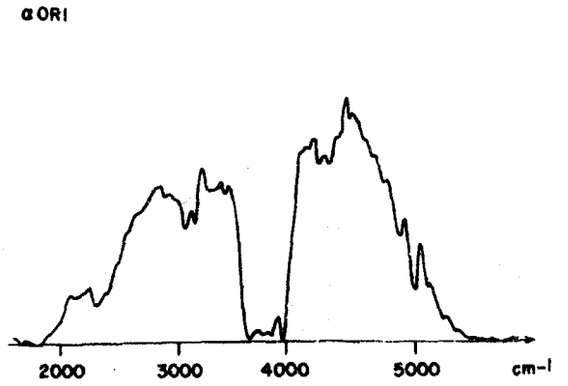


Figure 45-8. αORI Spectrum, Uncorrected for Sky Emission or Atmospheric Transmission

References

1. Girard, A. (1964) Principe d'un procédé d'analyse d'image par transformée de Fourier, *La Recherche Aérospatiale* **99**: 27.

2. Berny, J. (1965) Analyse d'image infrarouge unidimensionnelle par transformation de Fourier, *Note Technique ONERA* 85 (thesis).

3. Berny, J., Girard, A., and Margerand, D. (1967) Analyse d'image par transformation de Fourier en éclairage incohérent, *Optica Acta* **14**: 3, 267.

4. Bouchareine, P., and Connes, P. (1963) *J. Phys. Radium* **24**: 134.

5. Girard, A. (1970) *C. R. Ac. Sc. (Paris)* t. 270, B pp. 463–466.

Contents

46. Hadamard-Transform Multiplex Grating Spectrometer

John A. Decker, Jr.
Comstock and Wescott, Inc.
Cambridge, Massachusetts

Abstract

A series of Hadamard-transform multiplex grating spectrometers which operate on the principle of multi-slit optical encoding at the spectrometer exit focal plane, and which represent the first successful operation of a Hadamard-transform spectrometer are described.

46–1 INTRODUCTION

The idea of multiplex-encoding the optical output of a spectrometer, to overcome the luminosity limitation of a sequentially-scanned dispersive instrument[1,2] by allowing the photodetector to sense a multiplicity of optical wavelengths simultaneously, goes back to the original work of Golay[3] and Fellgett.[4,5] The possibility of realizing a multiplex grating spectrometer by means of optical coding based on Hadamard matrices was first pointed out by Ibbett, Aspinall and Grainger[6] and, shortly afterwards, independently by Decker and Harwit.[7] The mathematical properties of the Hadamard-transform codes have been described previously by Sloane, Fine, Phillips and Harwit[8] and by Pratt, Kane and Andrews,[9] who also described two-dimensional Hadamard-transform coding techniques. Harwit, Phillips, Sloane and Fine[10] recently extended this theoretical work to dispersive spectrometers using Hadamard-transform coding at both the entrance and exit focal planes in a manner analogous to Girard's original grille spectrograph.[11] Since several accompanying papers will cover the theoretical aspects of Hadamard-transform spectrometers for both the singly (that is, exit only) and doubly (that is, entrance and exit) encoded cases, this discussion will be limited to the description of Comstock and Wescott's recent experimental work with exit focal-plane encoded Hadamard-transform spectrometers.[12]

46–2 COMSTOCK AND WESCOTT PROTOTYPE SPECTROMETER

The original Comstock and Wescott prototype Hadamard-transform spectrometer (HTS) shown in (Figure 46-1) was constructed by modifying a 0.25-m Jarrell-Ash Ebert monochromator. This required fitting it with a cyclic 19-slot S^7-matrix exit focal-plane optical encoding mask[8] and the appropriate mask-drive system. This instrument was, in all probability, the first successfully-operated HTS. In its basic (that is, conventional monochromator) form, the instrument was equipped with 0.1-mm wide entrance and exit slits and a 295-groove/mm grating blazed for a wavelength of 2.1 μ; it had a resolution (limited by aberrations) of about 10 Å in second order.

The Hadamard-transform encoder, shown on Figure 46-2, consisted of a photoetched mask mounted at the spectrometer exit focal plane, containing a 37-slot cyclic array of 0.1-mm wide transparent and opaque slots of 19-slot periodicity,

0110000101011110011011000010101111001,

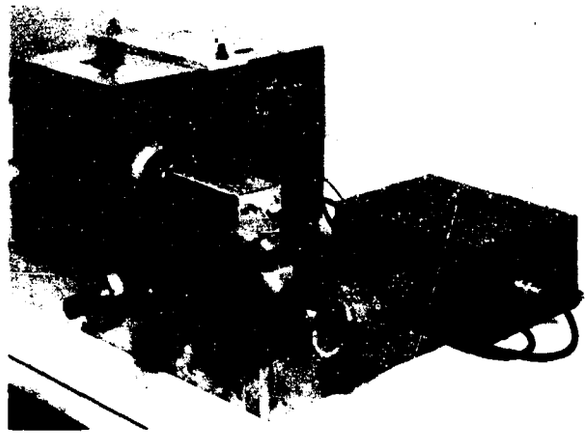


Figure 46-1. Comstock and Westcott Prototype Hadamard-Transform Spectrometer

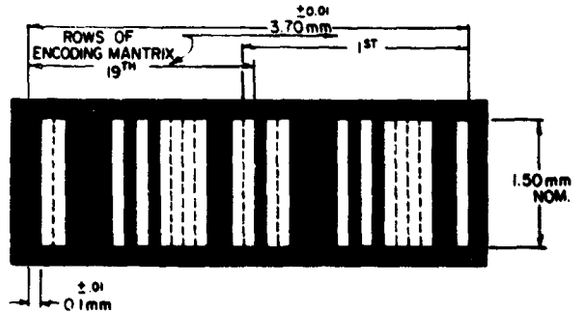


Figure 46-2. Comstock and Wescott Prototype 19-slot Hadamard-Transform Encoding Mask

(where 0 denotes opaque and 1 denotes transparent) derived from the S^T-matrix

$$[S_{19}{}^T] = \begin{bmatrix} 1011000010101111001 \\ 1101100001010111100 \\ 0110110000101011110 \\ 0011011000010101111 \\ 1001101100001010111 \\ 1100110110000101011 \\ 1110011011000010101 \\ 1111001101100001010 \\ 0111100110110000101 \\ 1011110011011000010 \\ 0101111001101100001 \\ 1010111100110110000 \\ 0101011110011011000 \\ 0010101111001101100 \\ 0001010111100110110 \\ 0000101011110011011 \\ 1000010101111001101 \\ 1100001010111100110 \\ 0110000101011110011 \end{bmatrix}$$

discussed by Sloane *et al.*[8] It provided 19 quasi-orthogonal mask patterns by moving to 19 successive positions in one-slot width (that is, 0.1 mm) steps along the plane of the mask, normal to the individual slots. A 1.90-mm wide field stop, which replaced the conventional exit slit, ensured that only 19 mask slots were exposed to the exiting radiation at any one time. The mask was moved from position to position by a manually driven micrometer screw; a second perpendicular micrometer screw provided a sensitive focus adjustment.

All radiation exiting the encoding mask was collected onto a single lead sulfide (PbS) infrared detector with an active area of 1 mm×3 mm. Signals from this detector were passed through a unity-gain FET preamplifier, a variable-gain signal amplifier and a synchronous demodulator gated by a signal from the 800-Hz tuning-fork chopper, then displayed on a digital voltmeter. The data were recorded manually, and were decoded, normalized, and plotted on a Telcomp II time-sharing computer system, using the straightforward but slow technique of multiplying the data by the inverse of the mask coefficient matrix.

A typical laboratory spectrum (of the mercury emission lines near 1.7 μ) obtained with the Comstock and Wescott Prototype Spectrometer is given in Figure 46-3(a) in comparison with a spectrum (Figure 46-3(b)) obtained using this optical system

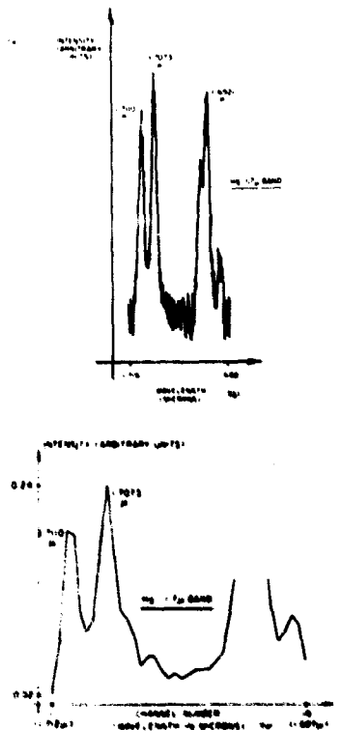


Figure 46-3. Comparison Spectra of the 1.7 μ Mercury Emission Lines: (a) As Obtained in the Hadamard-transform Mode; (b) As Obtained Using the Same Optical System as a Conventional Monochromator

as a conventional monochromator. Note that there are no detectable systematic errors and that the instrument's resolving power ($\lambda/\Delta\lambda \simeq 1000$ in first order) is limited only by the entrance slit width and the quality of the basic optical system: for example, the flat top of the line at 1.7110 μ is the image of the entrance slit as seen with the encoder fortuitously aligned with the line image. However, while the two spectra in Figure 46-3 were obtained in approximately the same total observing time, they are not to be interpreted as representative of the signal-to-noise ratio comparison between conventional spectrometers and HTS instruments. Since the noise level of an infrared photodetector is roughly proportional to its area, a strict signal-to-noise ratio comparison would further require the de-dispersion of the encoded radiation exiting the mask (for example, by a second reversed-pass through the spectrometer) to enable it to be brought to a focus on roughly the same size detector as would be optimally used by the system operating as a conventional monochromator. This is one of the design features of the AFCRL/Comstock and Wescott Demonstration Spectrometer discussed in the next section. The purpose of presenting the comparison spectrum here is to demonstrate the spectral fidelity of the HTS instrument and the freedom from systematic errors obtainable with even this somewhat crude prototype spectrometer.

46-3 COMSTOCK AND WESCOTT/AFCRL DEMONSTRATION SPECTROMETER

The maximum signal-to-noise ratio improvement to be expected for an HTS instrument using a 19-slot 1,0 code is only about a factor of 2.18, referenced to a monochromator[4,5,8] even when the exiting radiation is de-dispersed onto a minimum-sized detector. This is rather hard to verify experimentally, as it barely exceeds the minimum noticeable change in signal-to-noise ratio.[13] Comstock and Wescott therefore built, in a joint program with AFCRL, an HTS which used a 255-slot encoder. It has a theoretical maximum signal-to-noise ratio improvement factor of about 8.0, which is much easier to quantitatively verify by experiment. This instrument (shown on Figure 46-4) is being demonstrated at this conference. An optical schematic of this instrument is given as Figure 46-5. The basic dispersive system was built around a Minuteman Instrument Company Model 305 0.5-meter fully corrected Czerny-Turner spectrograph specifically designed for wide-exit focal-plane applications. It was equipped with a 0.005 to 3.0-mm wide bilaterally-variable entrance slit (normally used at a setting of 0.1 mm) and a 150-groove/mm grating blazed for a wavelength of 2.0 μ. In this configuration it had a first order resolution of about 10 Å (2.5 cm^{-1} at 2 μ) with a 0.1-mm wide entrance slit (that is, a resolving power of about 2000) and a well-focussed exit focal-plane width of about 10.2 cm.

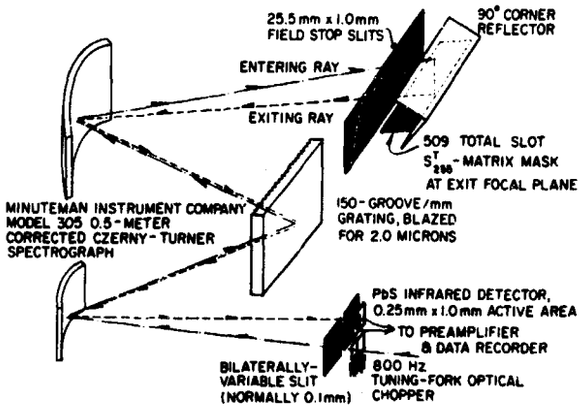


Figure 46-5. Optical Schematic of Comstock and Wescott/AFCRL HTS Spectrometer

The Hadamard-transform encoding mask used a 255-slot cyclic S-matrix code, 509 total mask slots, (Figure 46-6). Its code pattern,

0000 00101 10001 11101 00001 11111 11001
00001 01001 11110 10101 01110 00001 10001
01011 00110 01011 11110 11110 01101 11011
10010 10100 10100 01001 01101 00011 00111
00111 10001 10110 00010 00101 11010 11110
11011 11100 00110 10011 01011 01101 01000
00100 11101 10010 01001 10000 00111 01001
00011 10001 00000 00101 10001 11101 00001
11111 11001 00001 01001 11110 10101 01110
00001 10001 01011 00110 01011 11110 11110
01101 11011 10010 10100 10100 01001 01101
00011 00111 00111 10001 10110 00010 00101
11010 11110 11011 11100 00110 10011 01011
01101 01000 00100 11101 10010 01001 10000
00111 01001 00011 10001,

Figure 46-4. Comstock and Wescott/AFCRL 255-Slot Demonstration HTS

Figure 46-6. 255-Slot Encoding Mask for Comstock and Wescott/AFCRL HTS Spectrometer

was computed using the shift-register logic given by Baumer.[14] This mask is located at the exit focal plane of the spectrometer, mounted so that it bisects a 90° corner mirror which returns the encoded light beam back through the dispersive optics and shifts it from above to below the optical axis (see Figure 46-7).

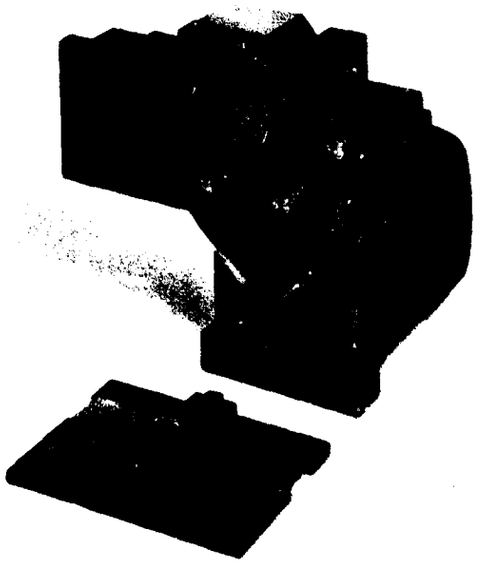


Figure 46-7. HTS Exit Focal Plane: Encoding Mask, Mask-drive Mechanism, Reversed-pass Corner Mirror, and Field Stop

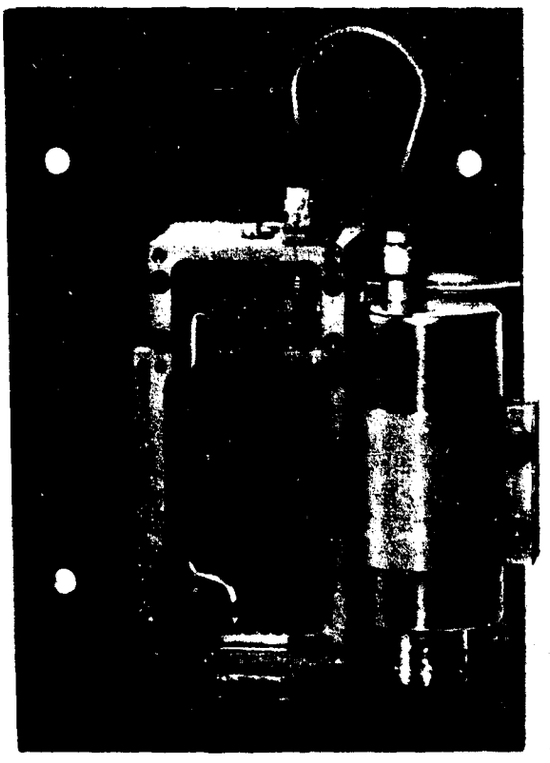


Figure 46-8. HTS Entrance Focal Plane: Tuning-fork Chopper, Detector Mounting and Preamplifier

This reversed pass de-disperses the encoded beam and allows it to be brought to a focus at the entrance focal plane on approximately the same size detector as would have been required using the optical system as a monochromator. As the entrance rays are inserted into the spectrometer slightly below the optical axis, the exiting rays emerge slightly above the optical axis, and are detected by a 0.25-mm×1.0-mm PbS cell mounted on the upper portion of the entrance slit, above the tuning-fork chopper (see Figure 46-8). The mask is moved sequentially by a stepping motor-driven belt drive (Figure 46-7) which is controlled by an electronic indexing circuit. The signals from the detector are amplified by a low-noise high-gain preamplifier, passed through an analog-to-digital convertor, and recorded in standard ASCII format on punched paper tape. The necessary sequencing and gating signals are provided by the indexing circuit that controls the mask-drive stepping motor. The data tapes are then read into a PDP-10 time-sharing computer for decoding and automatic plotting of the output spectra. The FORTRAN IV decoding program takes full account of the binary nature of the Hadamard coding matrix to minimize computer running time. The actual decoding computation, exclusive of the time needed to read in the data tapes and plot the spectrum, is approximately 1.5 sec for the 255×255 coding matrix.

The 255-slot instrument became fully operational only two days prior to its shipment to Aspen for the conference display; therefore no detailed analysis of its performance can be given at this time. However, a preliminary HTS scan of a 0.34-μ wide region of the mercury emission spectrum near 1.6 μ is given as Figure 46-9. Note that this spectrum was obtained from a set of Hadamard-encoded data points which had a total dynamic range of less than 52 percent of the peak signal (that is, for this specific spectrum the encoded data points varied from a maximum of 0.731 to a minimum of 0.240, in arbitrary units); this appears to be typical. No accurate test of the signal-to-noise gain in comparison with a monochromator has been made to date, but it will be reported when the work is complete.

A further modification of this basic instrument that will increase its signal-to-noise ratio gain to a factor of about 22.6, will double its resolution (to 1.25 cm^{-1} at 2.0 μ) and will, at the same time, increase its scanning speed to 2-to-10 spectral scans per second is currently nearing completion. This uses the cyclic 2047×2047 S-matrix to generate the encoding mask whose mask pattern is given on Figure 46-10 and

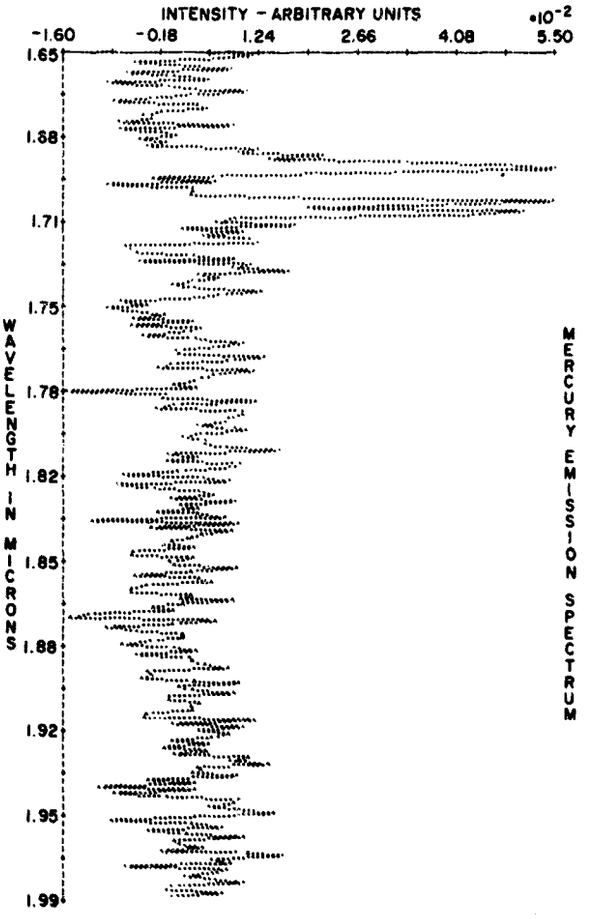


Figure 46-9. Preliminary Spectrum Obtained With the 255-slot HTS—a 0.34-μ Wide Region of the Mercury Emission Spectrum Near 1.6 μ

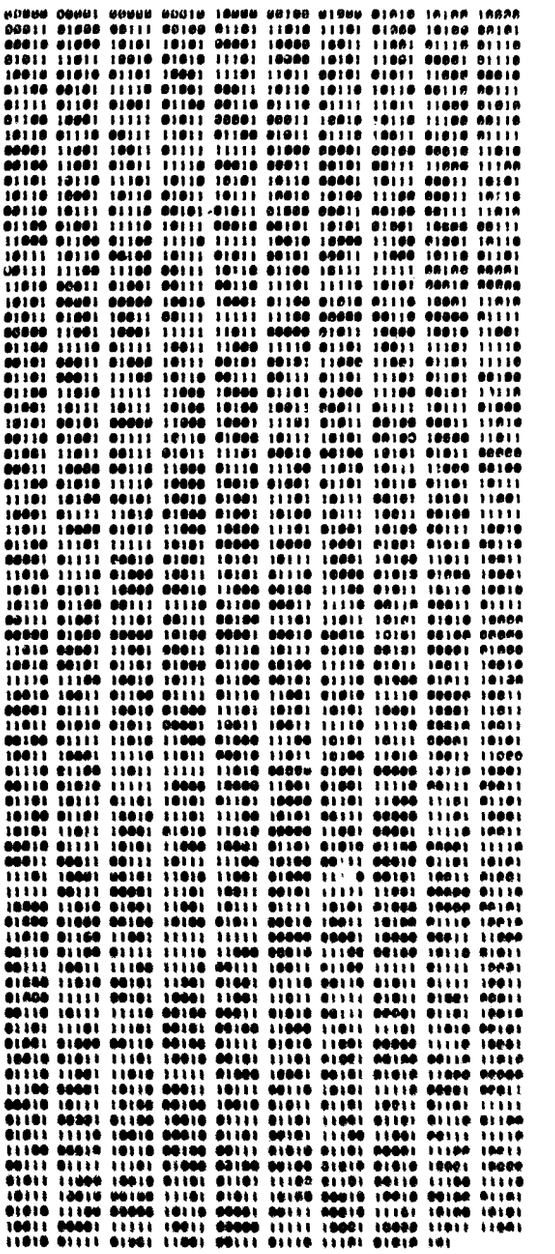


Figure 46-10. 4003-slot Encoding Mask Code Used With 2047-slot HTS System

it will contain a total of 4094 0.03-mm wide slots. This mask will be driven continuously (rather than stepwise sequentially), and will feed data to a magnetic-tape analog data recorder. The FORTRAN IV data decoding program will have a running time on the PDP-10 of only a fraction of a minute (without using the recently developed fast Hadamard-transform programs[9,13,16]) but may require batch processing to obtain reasonable input/output times. This HTS system should be completed within a few weeks, and should be fully operational by late spring.

46-4 COMPARISON WITH OTHER MULTIPLEX SYSTEMS

From the point of view of the experimentalist, the HTS is emerging as an instrument which displays the classical multiplex advantage,[4,5] yet uses the comparatively simple, reliable and thoroughly proven technology of dispersive spectrometers. Its throughput, limited by the entrance slit width, is somewhat less than that of interferometric instruments (although the proposed doubly-encoded HTS[8,10] should have approximately the interferometer's throughput) and its spectral resolution is limited by the grating size as well as by the entrance slit width. Additionally, it shares with nearly all multiplex spectrometers the need for a digital computer to transform the encoded data into a usable spectrum. However, it typically requires micrometric, rather than interferometric, tolerances in fabrication and alignment, and it can easily be modified to record only those wavelengths

of particular interest, which is difficult and/or inefficient to do with interferometric spectrometers. Given an existing dispersive spectrometer (or, more likely, spectrograph or polychromator) with the required off-axis image quality and exit focal-plane width, it also provides the possibility of achieving multiplex performance by means of a modification or retro-fit at a fraction of the cost of an interferometric instrument.

More fundamentally, however, the signal from an HTS lacks the large zero order spike which is characteristic of the more straightforward Fourier transform spectrometers,[17] and hence it has a substantially lower dynamic range. This is directly reflected in a lower required data transmission/recording bit-rate and/or bandwidth. In addition, the computer decoding of the data from an HTS system, using fast-Hadamard-transform programs, is probably up to an order of magnitude faster than the equivalent Fourier-transform inversion required by FTS systems.[9] In general, the HTS should eventually be preferred over interferometric multiplex systems for those applications where the total system performance is limited by bandwidth, maximum data-rate, computer time for data decoding, or mechanical reliability, and for observations of faint point-sources (for example, stars) where the interferometer's throughput advantage does not come into play.

References

1. Jacquinot, P. (1954) *J. Opt. Soc. Am.* **44**: 761.

2. Jacquinot, P. (1960) *Reports Prog. Phys.* **23**: 267.

3. Golay, M. J. E. (1949) *J. Opt. Soc. Am.* **39**: 437.

4. Fellgett, P. (1951) Ph.D. thesis, Cambridge University.

5. Fellgett, P. (1958) *J. Phys. Radium* **19**: 187.

6. Ibbett, R. N., Aspinall, D., and Grainger, J. F. (1968) *Appl. Opt.* **7**: 1089.

7. Decker, J. A., and Harwit, M. O. (1968) *Appl. Opt.* **7**: 2205.

8. Sloane, N. J. A., Fine, T., Phillips, G. P., and Harwit, M. (1969) *Appl. Opt.* **8**: 2103.

9. Pratt, W. K., Kane, J., and Andrews, H. C. (1969) *Proc. IEEE* **57**: 58.

10. Harwit, M., Phillips, P. G., Sloane, N. J. A., and Fine, T. (1970) *Appl. Opt.* to be published.

11. Girard, A. (1963) *Appl. Opt.* **2**: 79.

12. Decker, J. A., and Harwit, M. (1969) *Appl. Opt.* **8**: 2552.

13. Rolfe, J., and Moore, S. E. (1970) *Appl. Opt.* **9**: 63.

14. Baumert, L. D. (1964) in *Digital Communications with Space Applications*, S. W. Golomb, Ed., Prentice-Hall, Englewood Cliffs, New Jersey, pp. 17–32, 47–64 and 165–195.

15. Shanks, J. L. (1969) *IEEE Trans. Computers* **C18**: 457.

16. Gentleman, W. M. (1968) *Bell System Tech. J.* **47**: 1099.

17. Mertz, L. (1965) *Transformations in Optics*, John Wiley and Sons, New York.

Discussion

Q. (Robert Dick, Barringer Research, Ltd.): You speak of an instrument in which the total throughput is modulated. What limitations does this impose on measurement of very weak absorption spectra?

A. (John A. Decker): None. It may improve the measurement, in that the multiplex advantage would improve signal-to-noise over scanning measurements. We have not yet conducted the appropriate experiment with the "Doubly Multiplexed" systems (as in Sloane's and Phillips' paper), but we have detected weak spectral structure successfully with the "Singly Multiplexed" instrument I described.

Q. (Robert Dick): Have you tried to apply this to an Ebert configuration with large curved slits?

A. (John A. Decker): Curved slits are OK if the radius of curvature doesn't change too much over the exit field; a corrected Czerny-Turner designed specifically for wide-field, high-slit use (a Minuteman 305, for one commercially available example) is better, and is what we have used so far—with straight slits. A "Doubly Multiplexed" system cannot, in general, use the simple Ebert mounting, because of lateral magnification problems. A Littrow is better in that case, and is what Perry Phillips used.

Contents

47. Reducing the Number of Measurements in Doubly Multiplexed Spectrometers

N. J. A. Sloane
Bell Telephone Laboratories, Incorporated
Murray Hill, New Jersey

Abstract

A doubly multiplexed grating spectrometer is described having N entrance slots and N exit slots. The simplest operation of this device requires N^2 observations to estimate $2N-1$ unknown spectral elements, but it is shown here that $2N-1$ observations always suffice. An algorithm is given for determining which $2N-1$ measurements should be made, and for forming the estimate of the spectrum from these measurements. Numerical results are presented for the values $N=3$, 7, 11, and 15, in the case that both entrance and exit masks are obtained from cyclic Hadamard matrices. (Other values of N have been analysed, but space does not permit their inclusion here.)

47–1 INTRODUCTION

A summary of this paper is given in the abstract. This paper is a sequel to the papers by Sloane et al.[1] and Harwit et al.[2] but can be read independently of them. For a comparison of these methods with other spectrometric techniques, and for experimental results, see the papers of Decker and Phillips (Chapters 46 and 48).

47–2 DESCRIPTION OF A DOUBLY MULTIPLEXED SPECTROMETER

47–2.1 Encoding the Input to the Grating

An encoding mask is placed at the position of the entrance slit of a conventional grating spectrometer. Instead of passing light through just one entrance slot, this mask is N slot widths wide. In any given mask position, light from the source is permitted to pass through about half of these N slots and is blocked from passage through the others. In general, there are N different mask positions in which different combinations of slots are permitted to pass light from the source into the spectrometer. The encoding process then consists of the successive use of each of the N mask positions to pass light through N different preselected combinations of entrance slot locations.

47–2.2 Encoding the Output

A second encoding mask is placed in the exit focal plane of the spectrometer. This mask has N different positions, each position passing light through some of the N exit slot locations in a preselected way. By making measurements of the intensity of radiation passing through different combinations of entrance

and exit mask positions, the radiation spectrum can be recovered.

47-3 ANALYSIS

Let $\epsilon=(\epsilon_{ir})$ be the $N \times N$ matrix describing the entrance mask, where $\epsilon_{ir}=1$ or 0, according as the rth entrance slot is open or closed when the entrance mask is in position i, $(1\leq i\leq N,\ 1\leq r\leq N)$. Similarly let $\chi=(\chi_{ij})$ describe the exit mask.

47-3.1 The Basic Equation

When the entrance mask is in position i and the exit mask is in position j, the detector measures

$$\eta_{i,j}=\sum_{r=1}^{N}\sum_{s=1}^{N}\epsilon_{i,r}\psi_{r,s}\chi_{j,s}+\nu_{i,j}, \qquad (47\text{-}1)$$

where $\psi_{r,s}$ is the measurement that would be made by a noiseless detector placed at the sth exit slot when signal enters only through the rth entrance slot, and $\nu_{i,j}$ is the detector noise for measurement (i,j). Assume that the optical source is spatially homogeneous, and that the spectrometer optics are arranged so that the spectrum produced by the rth entrance slot acting alone is a shift by r places of the spectrum produced by the first entrance slot acting alone. Thus, there are $2N-1$ unknown spectral components:

$$\psi_{-(N-1)},\ \ldots,\ \psi_{-1},\ \psi_0,\ \psi_1,\ \ldots,\ \psi_{N-1}, \qquad (47\text{-}2)$$

given by $\psi_{r,s}\triangleq\psi_{r-s}$. Then Eq. (47-1) may be rewritten in matrix form as

$$\eta=\epsilon\mathcal{J}\chi^T+\nu, \qquad (47\text{-}3)$$

where $\eta=(\eta_{i,j})$, $\nu=(\nu_{i,j})$, and

$$\mathcal{J}=\begin{pmatrix}\psi_0 & \psi_{-1} & \cdots & \psi_{-N+1}\\ \psi_1 & \psi_0 & \cdots & \psi_{-N+2}\\ \cdots & \cdots & \cdots & \cdots\\ \psi_{N-1} & \psi_{N-2} & \cdots & \psi_0\end{pmatrix}. \qquad (47\text{-}4)$$

Equations (47-1) and (47-3) are equivalent forms of the basic equation of the spectrometer.

The detector noise components $\nu_{i,j}$ are assumed to have average value zero and to be uncorrelated.

47-4 THE ESTIMATION PROBLEM

There is the possibility of making up to N^2 measurements $\eta_{i,j}$, $(i,j=1,\ldots,N)$ to estimate $2N-1$ unknown ψ_t's. Given these measurements, the unknown ψ_t's must be estimated. As justified by Sloane *et al.*[1] for a singly multiplexed spectrometer, only linear unbiased estimators will be considered.

One method is to make the full set of N^2 measurements. This case has been considered in detail by Harwit *et al.*[2] The next section shows that the number of measurements needed can be greatly reduced, from N^2 to $2N-1$.

47-5 USING ONLY 2N−1 MEASUREMENTS

Provided the mask matrices ϵ and χ are nonsingular, it is possible to obtain linear unbiased estimates of the $2N-1$ unknown ψ_t's by measuring only $2N-1$ $\eta_{i,j}$'s. This may be shown as follows. From the basic Eq. (47-3), the expected values $E(\eta_{i,j})\triangleq\bar{\eta}_{i,j}$ and the unknown ψ_t are related by the equation

$$\begin{pmatrix}\bar{\eta}_{11} & \bar{\eta}_{12} & \cdots & \bar{\eta}_{1N}\\ \bar{\eta}_{21} & \bar{\eta}_{22} & \cdots & \bar{\eta}_{NN}\\ & \cdots & & \\ \bar{\eta}_{N1} & \bar{\eta}_{N2} & \cdots & \bar{\eta}_{NN}\end{pmatrix}=$$

$$(\epsilon)\begin{pmatrix}\psi_0 & \psi_{-1} & \cdots & \psi_{-N+1}\\ \psi_1 & \psi_0 & \cdots & \psi_{-N+2}\\ & \cdots & & \\ \psi_{N-1} & \psi_{N-2} & \cdots & \psi_0\end{pmatrix}(\chi^T). \qquad (47\text{-}5)$$

If ϵ and χ are nonsingular, this transformation is invertible, and there exists a $(2N-1)\times N^2$ matrix γ such that

$$\begin{pmatrix}\psi_{N-1}\\ \vdots\\ \psi_0\\ \vdots\\ \psi_{-N+1}\end{pmatrix}=(\gamma)\begin{pmatrix}\bar{\eta}_{11}\\ \bar{\eta}_{12}\\ \vdots\\ \bar{\eta}_{1N}\\ \bar{\eta}_{21}\\ \vdots\\ \bar{\eta}_{NN}\end{pmatrix}, \qquad (47\text{-}6)$$

where γ has rank $2N-1$. Then γ must contain $2N-1$ linearly independent columns, corresponding say to the coordinates

$$(i_1,j_1),\ (i_2,j_2),\ \ldots,\ (i_{2N-1},j_{2N-1}) \qquad (47\text{-}7)$$

of the $\bar{\eta}$ vector. Expressing all columns of γ in terms of these $2N-1$ independent columns, it follows that

there is a matrix $G=(g_{ij})$ such that

$$\begin{pmatrix} \psi_{N-1} \\ \vdots \\ \psi_0 \\ \vdots \\ \psi_{-N+1} \end{pmatrix} = (G) \begin{pmatrix} \eta_{i_1 j_1} \\ \eta_{i_2 j_2} \\ \vdots \\ \eta_{i_{2N-1} j_{2N-1}} \end{pmatrix}. \qquad (47\text{-}8)$$

Then from Eq. (47-8), if measurements

$$\eta_{i_1 j_1}, \eta_{i_2 j_2}, \cdots, \eta_{i_{2N-1} j_{2N-1}}, \qquad (47\text{-}9)$$

are made, and the matrix $G=(g_{ij})$ is known, linear unbiased estimators for the unknown ψ_t's are given by

$$\hat{\psi}_t = \sum_{r=1}^{2N-1} g_{tr} \eta_{i_r j_r}. \qquad (47\text{-}10)$$

(In general, there may be many possible choices for the set (47-9) of independent measurements and for the matrix G.)

47–6 AN EXAMPLE

Suppose $N=2$, and that ϵ and χ are equal to the 2×2 Hadamard matrix $\begin{pmatrix} +1 & +1 \\ +1 & -1 \end{pmatrix}$. (As pointed out by Sloane *et al.*,[1] mask weights of ± 1 can be obtained by chopping between transmitted and reflected radiation.) In this case, Eq. (47-5) becomes

$$\begin{pmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \end{pmatrix} = \begin{pmatrix} \psi_1+2\psi_0+\psi_{-1} & \psi_1-\psi_{-1} \\ -\psi_1+\psi_{-1} & -\psi_1+2\psi_0-\psi_{-1} \end{pmatrix}. \qquad (47\text{-}11)$$

There are many ways of inverting this transformation; one is given by

$$\begin{pmatrix} \psi_1 \\ \psi_0 \\ \psi_{-1} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & 0 & -\frac{1}{4} \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{1}{2} & -\frac{1}{4} \end{pmatrix} \begin{pmatrix} \eta_{11} \\ \eta_{12} \\ \eta_{21} \\ \eta_{22} \end{pmatrix}. \qquad (47\text{-}12)$$

The first, third, and fourth columns of this matrix are independent, and Eq. (47-8) becomes

$$\begin{pmatrix} \psi_1 \\ \psi_0 \\ \psi_{-1} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & -\frac{1}{2} & -\frac{1}{4} \\ \frac{1}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \end{pmatrix} \begin{pmatrix} \eta_{11} \\ \eta_{21} \\ \eta_{22} \end{pmatrix}. \qquad (47\text{-}13)$$

Thus, $2N-1=3$ measurements η_{11}, η_{12}, η_{22} should be made, and the ψ_t's are estimated by

$$\begin{aligned} \hat{\psi}_1 &= \tfrac{1}{4}\eta_{11} - \tfrac{1}{2}\eta_{21} - \tfrac{1}{4}\eta_{22}, \\ \hat{\psi}_0 &= \tfrac{1}{4}\eta_{11} + \tfrac{1}{4}\eta_{22}, \\ \hat{\psi}_{-1} &= \tfrac{1}{4}\eta_{11} + \tfrac{1}{2}\eta_{21} - \tfrac{1}{4}\eta_{22}. \end{aligned} \qquad (47\text{-}14)$$

47–7 NUMERICAL RESULTS

In general, there does not seem to be any simple way of determining which set (47-9) of measurements to make nor of finding the matrix G to use in Eq. (47-10) for estimating the spectrum.

Table 47-1 presents some numerical results obtained with the aid of a computer, for the values $N=3$, 7, 11, and 15. (Corresponding results for $N=19$, 23, 31, and 35 have been obtained, but space does not permit their inclusion here.) For each N are given (*a*) the value of N, (*b*) the cyclic matrix S of mask weights used for both ϵ and χ, (*c*) a possible set of $2N-1$ independent measurements (47-9), and finally (*d*) the corresponding matrix G to be used in Eq. (47-10) for estimating the unknown spectral components ψ_t. The matrix of mask weights used here for both ϵ and χ is the matrix S defined in Sloane *et al.*[1] and Harwit *et al.*[2] Let H be a normalized $(N+1) \times (N+1)$ Hadamard matrix of $+1$'s and -1's with the first row and column consisting of $+1$'s. Then S is obtained by deleting the first row and column of H and replacing $+1$'s by 0's and -1's by 1's. [Incidentally, $S^{-1} = \frac{2}{N+1}(2S^T - J)$, where I is a unit matrix and J is an $N \times N$ matrix of $+1$'s.] The cyclic versions of S used here were obtained from Baumert.[3]

The computer program used to obtain these results proceeded as follows. First, the mask matrices ϵ and χ were read in. Second, the matrix product on the right of Eq. (47-5) was evaluated. Third, a search was made for $2N-1$ independent measurements $\eta_{i_r j_r}$. (This search must always be successful, as proved.) Fourth, the matrix equation relating the independent $\eta_{i_r j_r}$ to the ψ_t was obtained. Fifth and last, this equation was inverted to give the matrix G.

Table 47-1. Sets of Independent Measurements and Coefficients of Linear Unbiased Estimators

N = 5

BOTH MASKS ETA AND CHI ARE EQUAL TO

101
110
011

2N-1 INDEPENDENT MEASUREMENTS ARE
(READ DOWN - TOP VALUE IS I COORDINATE,
BOTTOM VALUE IS J COORDINATE)

1	1	1	2	2
1	2	3	2	3

G MATRIX FOR USE IN EQUATION (8) IS

COLUMN	1	2	3	4	5
ROW					
1	0.2500	0.7500	-0.2500	-0.2500	0.
2	-0.2500	0.2500	0.7500	0.0000	-0.5000
3	0.2500	-0.2500	-0.2500	0.5000	0.
4	-0.2500	0.2500	-0.2500	0.0000	0.5000
5	0.2500	-0.2500	0.7500	-0.5000	0.

N = 7

BOTH MASKS ETA AND CHI ARE EQUAL TO

0101110
0010111
1001011
1100101
1110010
0111001
1011100

2N-1 INDEPENDENT MEASUREMENTS ARE
(READ DOWN - TOP VALUE IS I COORDINATE,
BOTTOM VALUE IS J COORDINATE)

1	1	1	1	1	1	1	2	3	3	3	3	3
1	2	3	4	5	6	7	3	1	2	3	4	5

G MATRIX FOR USE IN EQUATION (8) IS

COLUMN	1	2	3	4	5	6	7
ROW							
1	-0.3125	-0.2500	-0.1250	-0.0625	-0.0625	0.0625	-0.2500
2	-0.0625	-0.0000	0.3750	-0.0625	-0.0625	-0.1875	0.2500
3	0.1875	-0.2500	-0.1250	0.4375	-0.0625	0.0625	-0.2500
4	-0.3125	-0.0000	0.1250	-0.0625	0.4375	0.0625	0.0000
5	-0.0625	-0.2500	0.1250	-0.0625	-0.0625	0.3125	0.0000
6	-0.0625	0.0000	-0.1250	-0.0625	-0.0625	-0.1875	0.2500
7	0.4375	0.2500	-0.3750	-0.0625	-0.0625	-0.1875	0.0000
8	0.1875	0.5000	-0.3750	-0.0625	-0.0625	0.0625	-0.0000
9	-0.0625	-0.2500	0.6250	-0.0625	-0.0625	0.3125	0.0000
10	-0.3125	-0.5000	0.6250	0.4375	-0.0625	0.0625	-0.0000
11	-0.3125	-0.2500	0.3750	-0.0625	0.4375	-0.4375	-0.2500
12	-0.0625	0.5000	-0.6250	-0.0625	-0.0625	-0.1875	-0.2500
13	-0.0625	-0.2500	0.1250	-0.0625	-0.0625	0.3125	0.0000

COLUMN	8	9	10	11	12	13
ROW						
1	1.1250	-0.1250	0.1875	0.2500	-0.4375	0.0625
2	-0.3750	0.1250	-0.3125	0.0000	0.3125	0.0625
3	0.1250	-0.1250	0.1875	-0.2500	0.0625	0.0625
4	0.1250	-0.1250	-0.0625	0.2500	-0.1875	-0.1875
5	0.1250	0.1250	-0.0625	0.0000	0.0625	-0.1875
6	0.1250	0.1250	0.1875	-0.0000	-0.1875	0.0625
7	-0.3750	0.1250	-0.0625	-0.0000	0.0625	0.3125
8	-0.3750	-0.1250	-0.0625	-0.2500	0.3125	0.3125
9	0.1250	-0.3750	-0.0625	0.0000	0.0625	-0.1875
10	0.6250	-0.1250	-0.0625	0.2500	-0.1875	-0.6875
11	0.6250	0.3750	0.1875	0.2500	-0.4375	-0.4375
12	-0.3750	0.6250	0.1875	-0.0000	-0.1875	0.5625
13	-0.3750	-0.3750	0.4375	0.0000	0.5625	-0.1875

N = 11

BOTH MASKS ETA AND CHI ARE EQUAL TO

11021100010
01201110001
10110111000
01011011100
00101101110
00010110111
10001011011
11000101101
11100010110
01110001011
10111000101

2N-1 INDEPENDENT MEASUREMENTS ARE
(READ DOWN - TOP VALUE IS I COORDINATE,
BOTTOM VALUE IS J COORDINATE)

1	1	1	1	1	1	1	1	1	2	3	3	3	3	3	3	3
1	2	3	4	5	6	7	8	9	10	11	1	2	3	4	5	6	7	8	9

G MATRIX FOR USE IN EQUATION (8) IS

COLUMN	1	2	3	4	5	6	7
ROW							
1	-0.1389	0.2500	-0.1944	-0.2500	-0.2500	0.0278	0.1944
2	0.2500	-0.1389	0.3750	-0.0417	-0.0417	-0.1667	0.2083
3	0.1944	0.2500	-0.0278	0.4722	0.0278	-0.0833	-0.0833
4	-0.0833	0.1944	0.2083	-0.0972	0.3472	-0.0556	-0.1806
5	0.0833	-0.0833	0.1111	0.1111	-0.2222	0.3056	-0.3472
6	0.0278	0.0833	-0.2528	0.0417	0.0417	-0.2222	0.2361
7	-0.0278	0.0278	0.0556	-0.1667	0.0556	0.0833	-0.2222
8	-0.0833	-0.0278	-0.0694	-0.0417	-0.2639	0.0556	-0.0139
9	-0.0278	-0.0833	-0.0000	0.0000	0.1111	-0.1389	0.1667
10	-0.0833	-0.0278	0.0139	0.0972	0.0972	0.1111	-0.0417
11	0.0833	-0.0833	-0.0556	-0.0556	-0.0556	-0.0278	-0.0000
12	0.0278	0.0833	-0.0694	-0.0417	-0.0417	-0.0556	-0.0139
13	0.0833	0.0278	0.0833	-0.0833	-0.0833	-0.0833	-0.0833
14	0.0278	0.0833	0.0139	0.0972	-0.0139	0.0000	-0.0417
15	0.0833	0.0278	0.0833	0.0278	0.1389	0.0278	0.0278
16	-0.0833	0.0833	0.0694	0.1528	0.1528	0.2222	0.1250
17	-0.1389	-0.0833	0.1389	0.0833	0.0833	0.0278	0.1944
18	-0.1944	-0.1389	-0.1528	0.0417	-0.0694	-0.0000	-0.0972
19	0.0833	-0.1944	-0.1944	-0.2500	-0.1389	-0.1944	-0.1389
20	0.1389	0.0833	-0.2083	-0.1250	-0.0139	0.1111	-0.0417
21	0.0833	0.1389	0.0556	-0.2778	-0.2778	-0.1389	-0.0000

COLUMN	8	9	10	11	12	13	14
ROW							
1	-0.0278	-0.0278	0.1944	-0.8056	1.0000	-0.2222	-0.2222
2	-0.0278	-0.0278	-0.0139	0.2222	0.	-0.0139	-0.2500
3	-0.0278	-0.0278	0.1389	-0.0278	0.	-0.1667	-0.0000
4	-0.0278	-0.0278	0.0417	0.1111	0.	-0.0694	-0.1389
5	-0.0278	-0.0278	-0.0556	0.0278	0.	0.0278	-0.0556
6	-0.0278	-0.0278	-0.0972	-0.0556	0.	0.0694	0.0278
7	0.3056	-0.0278	-0.1111	-0.0833	0.	0.0833	0.0556
8	-0.0278	0.3056	-0.1250	-0.1111	0.	0.0972	0.0833
9	-0.0278	-0.0278	0.1667	-0.0833	0.	0.1389	0.0556
10	-0.0278	-0.0278	0.0694	0.1667	0.	-0.0972	0.1389
11	-0.0278	-0.0278	0.1111	0.0278	0.	-0.1389	-0.0556
12	-0.0278	-0.0278	-0.0139	0.1111	0.	-0.0139	-0.1389
13	-0.0278	-0.0278	0.0278	-0.0278	0.	-0.0556	0.0000
14	-0.0278	-0.0278	-0.1528	0.0556	0.	0.1250	-0.0833
15	-0.0278	-0.0278	-0.0833	-0.1389	0.	0.0556	0.1111
16	-0.0278	-0.0278	-0.0972	-0.0556	0.	0.0694	0.0278
17	-0.0278	-0.0278	0.1944	-0.1389	0.	-0.2222	0.1111
18	0.3056	-0.0278	0.0139	0.1667	0.	-0.0417	-0.1944
19	-0.0278	0.3056	0.0833	-0.0278	0.	-0.1111	-0.0000
20	-0.0278	-0.0278	-0.0417	0.1667	0.	0.3472	-0.1944
21	-0.0278	-0.0278	0.1111	-0.0833	0.	-0.1389	0.3889

COLUMN	15	16	17	18	19	20	21
ROW							
1	0.1111	0.0556	0.1667	0.2222	0.2222	-0.0556	-0.2222
2	-0.2778	0.1111	-0.0694	0.0139	0.0139	0.1389	-0.2917
3	-0.2222	-0.2778	-0.0000	-0.1667	-0.0556	0.0556	0.0139
4	0.0556	-0.2222	-0.2361	0.0694	-0.0417	0.0278	0.1528
5	-0.1111	0.0556	-0.1389	-0.1389	0.1944	0.0000	-0.1389
6	-0.0556	-0.1111	0.1250	-0.0694	-0.0694	0.1944	0.0694
7	0.0000	-0.0556	-0.0833	0.1389	-0.0833	-0.1111	0.1944
8	0.0556	0.0000	0.0417	0.0139	0.2361	-0.0833	-0.0139
9	0.0000	0.0556	-0.0278	-0.0278	-0.1389	0.1111	-0.1944
10	0.0556	-0.0000	-0.0417	-0.1250	-0.1250	-0.1389	0.0139
11	0.2222	0.0556	0.0278	0.0278	0.0278	-0.0000	-0.0278
12	-0.0556	0.2222	0.0417	0.0139	0.0139	0.0278	-0.0139
13	-0.1111	-0.0556	0.2222	0.0556	0.0556	0.0556	0.0556
14	-0.0556	-0.1111	-0.0417	0.2083	-0.0139	-0.0278	0.0139
15	-0.1111	-0.0556	-0.1111	-0.0556	0.1667	-0.0556	-0.0556
16	0.0556	-0.1111	-0.0972	-0.1806	-0.1806	0.0833	-0.1528
17	0.1111	0.0556	-0.1667	-0.1111	-0.1111	-0.0556	0.1111
18	0.1667	0.1111	0.1250	-0.0694	0.0417	-0.0278	0.0694
19	-0.1111	0.1667	0.1667	0.2222	0.1111	0.1667	0.1111
20	-0.1667	-0.1111	0.1806	0.0972	-0.0139	-0.1389	0.0139
21	-0.1111	-0.1667	-0.0833	0.2500	0.2500	0.1111	-0.0278

Table 47-1. Sets of Independent Measurements and Coefficients of Linear Unbiased Estimators (contd)

N = 15

BOTH MASKS ETA AND CHI ARE EQUAL TO

1000100110101111
1100010011010111
1110001001101011
1111000100110101
0111100010011010
1011110001001101
0101111000100110
0010111100010011
1001011110001001
0100101101111000
0010011010111100
0001001101011110
0000100110101111

2N-1 INDEPENDENT MEASUREMENTS ARE
(READ DOWN - TOP VALUE IS I COORDINATE,
BOTTOM VALUE IS J COORDINATE)

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

G MATRIX FOR USE IN EQUATION (8) IS

COLUMN

ROW	1	2	3	4	5	6	7
1	0.2969	0.1406	-0.0469	-0.0156	-0.0781	0.0781	-0.2031
2	-0.0156	0.0156	0.0781	-0.0156	0.1094	-0.2344	-0.1406
3	-0.3201	0.3906	0.4531	0.1719	-0.2656	0.2656	0.3594
4	0.0469	-0.3594	-0.1719	0.2344	0.2969	-0.1719	-0.2031
5	0.2344	-0.1719	-0.2344	-0.0781	0.2969	0.1406	-0.2031
6	-0.1406	0.2656	0.0781	-0.1406	-0.1406	0.2656	0.3594
7	-0.2031	0.0781	0.1406	-0.0156	-0.2031	0.0156	0.2969
8	0.1094	-0.2344	-0.1719	0.0469	0.0469	-0.1719	-0.2031
9	0.1094	0.0156	-0.0469	-0.0781	0.0469	-0.0469	-0.0781
10	-0.0156	-0.0469	0.0156	-0.0156	-0.0156	-0.0469	-0.1406
11	-0.0781	-0.0469	0.0156	0.0469	-0.0156	-0.0469	-0.0156
12	0.0469	-0.2344	-0.0469	0.0469	0.1094	-0.1094	-0.1406
13	0.0469	-0.1719	-0.1094	0.0469	0.1094	-0.0469	-0.1406
14	-0.2031	0.2031	0.1406	-0.0156	-0.0781	0.1406	0.2969
15	0.1094	-0.0469	-0.1094	-0.0156	-0.0156	0.0781	-0.0156
16	0.0469	0.2656	-0.0469	-0.1406	-0.0781	0.0781	0.1719
17	-0.0156	0.1406	0.0781	-0.1406	-0.1406	0.0156	-0.0156
18	-0.0781	0.2656	0.2031	0.0469	-0.2656	0.0156	0.2344
19	0.0469	0.0156	0.0781	0.1094	0.0469	-0.1719	-0.0781
20	-0.0156	0.2031	0.0156	0.0469	0.0469	0.1406	-0.0781
21	0.1094	-0.1094	0.0781	-0.0156	0.1094	0.0156	-0.0156
22	0.0469	-0.0469	-0.1094	0.1094	0.0469	0.0156	-0.0781
23	-0.1406	0.1406	0.0781	-0.0781	0.0469	0.0781	0.1719
24	0.1094	-0.3594	-0.0469	0.0469	0.0469	-0.0469	-0.2031
25	-0.0156	0.0781	0.0156	0.1094	-0.0156	-0.0469	0.2344
26	-0.0781	-0.1719	-0.2344	-0.0781	0.2344	-0.0469	-0.3906
27	0.2969	-0.6094	-0.0469	-0.0781	0.1094	-0.1094	-0.2656
28	0.0469	-0.2969	-0.3594	0.1719	0.1094	-0.2969	-0.2656
29	-0.2031	-0.1719	0.1406	-0.1406	0.1719	-0.1094	-0.0781

COLUMN

ROW	8	9	10	11	12	13	14
1	0.0469	0.0469	-0.1094	-0.1094	0.2031	-0.1094	-0.0156
2	-0.1406	-0.0781	0.0781	-0.1094	-0.1094	0.3281	-0.0156
3	0.0469	0.1094	0.0156	0.0781	-0.1094	-0.3594	-0.0156
4	-0.0781	-0.0781	-0.1094	0.0156	0.0781	0.0156	-0.0156
5	-0.0156	-0.1406	0.0156	-0.1094	0.0156	0.1406	-0.0156
6	-0.0156	0.0469	-0.0469	0.0156	-0.1094	-0.0469	-0.0156
7	0.1719	0.0469	-0.0469	-0.0469	0.0156	-0.1719	-0.0156
8	0.1094	0.1094	-0.0469	-0.0469	-0.0469	0.0781	-0.0156
9	-0.0156	0.1094	0.2031	-0.0469	-0.0469	-0.0469	-0.0156
10	-0.0156	-0.0781	0.1406	0.2031	-0.0469	0.0156	-0.0156
11	-0.0781	-0.0156	-0.0469	0.1406	0.2031	-0.0469	-0.0156
12	0.0469	-0.1406	0.0156	-0.0469	0.1406	0.2656	-0.0156
13	0.0469	-0.0156	-0.0469	0.0156	-0.0469	0.2031	0.2344
14	0.0469	0.1719	0.0781	-0.0469	0.0156	-0.1719	-0.0156
15	-0.0156	0.0469	0.0156	0.0781	-0.0469	0.0156	-0.0156
16	-0.0781	0.0469	0.0156	0.0156	0.0781	-0.1094	-0.0156
17	-0.0156	-0.0781	-0.0469	0.0156	0.0156	0.0781	-0.0156
18	-0.0781	0.1094	-0.1094	-0.0469	0.0156	-0.1094	-0.0156
19	0.0469	-0.0781	0.0156	-0.1094	-0.0469	0.0156	-0.0156
20	-0.1406	0.1094	-0.1094	0.0156	-0.1094	-0.1094	-0.0156
21	-0.1406	-0.2031	0.0781	-0.1094	0.0156	-0.0469	-0.0156
22	0.0469	-0.2031	-0.1719	0.0781	-0.1094	0.0781	-0.0156
23	-0.0156	0.1094	-0.1719	-0.1719	0.0781	-0.1719	-0.0156
24	0.1094	-0.1406	0.0781	-0.1719	-0.1719	0.2031	-0.0156
25	0.1094	0.1719	0.0156	0.0781	-0.1719	-0.2344	-0.0156
26	0.0469	-0.0156	0.0781	0.0156	0.0781	-0.0469	-0.0156
27	-0.0781	-0.1406	0.1406	0.0781	0.0156	0.2656	-0.0156
28	0.1719	-0.2656	0.0781	0.1406	0.0781	0.2031	0.2344
29	0.1719	0.1719	-0.0469	0.0781	0.1406	0.0781	-0.0156

COLUMN

ROW	15	16	17	18	19	20	21
1	-0.1719	0.1563	-0.3125	0.0937	0.0313	-0.0000	0.0625
2	-0.1719	0.1562	0.0000	-0.0312	0.1562	-0.0000	-0.1250
3	0.1406	-0.1563	0.3125	-0.4062	-0.4687	0.0625	0.2500
4	0.0781	-0.0937	-0.0625	0.3437	0.1562	-0.2500	-0.0625
5	-0.1719	0.1562	-0.2500	0.1562	0.2187	0.0625	-0.3125
6	-0.0469	0.0312	0.1250	-0.2812	-0.0938	0.1250	0.1250
7	0.1406	-0.1563	0.1875	-0.0937	-0.1563	0.0000	0.1875
8	0.0156	-0.0312	-0.1250	0.2187	0.1562	-0.0625	-0.0625
9	-0.1094	0.0938	-0.1250	-0.0312	0.0313	0.0625	-0.0625
10	-0.1094	0.0938	-0.0000	0.0312	-0.0312	0.0000	-0.0000
11	-0.0469	0.0313	0.0625	0.0312	-0.0313	-0.0625	0.0000
12	-0.1094	0.0938	-0.0625	0.2188	0.0313	-0.0625	-0.1250
13	-0.1719	0.1563	-0.0625	0.1562	0.0937	-0.0625	-0.1250
14	0.2656	-0.0312	0.1875	-0.2187	-0.1563	-0.0000	0.0625
15	0.1406	-0.1562	0.1250	0.0313	0.0938	0.0000	-0.0000
16	0.0781	-0.0938	-0.0625	-0.0312	0.0312	0.1250	0.0625
17	0.0781	-0.0938	-0.0000	-0.1562	0.1563	0.1250	0.1250
18	0.1406	-0.1563	0.0625	-0.2812	-0.2187	0.1875	0.2500
19	0.0781	-0.0937	-0.0625	-0.0312	-0.0937	-0.1250	0.1875
20	0.0781	-0.0937	0.0000	-0.2187	-0.0312	-0.0625	-0.0625
21	-0.0469	0.0313	-0.1250	0.0938	-0.0937	0.0000	-0.1250
22	-0.1094	0.0938	-0.0625	0.0312	0.0937	-0.1250	-0.0625
23	0.0156	-0.0313	0.1250	-0.1562	-0.0937	0.0625	-0.0625
24	-0.1094	0.0938	-0.1250	0.3437	0.0312	-0.0625	-0.0625
25	-0.1094	0.0937	0.0000	-0.0938	-0.0313	-0.1250	0.0000
26	-0.0469	0.0313	0.0625	0.1563	0.2187	0.0625	-0.2500
27	-0.3594	0.3438	-0.3125	0.5937	0.0313	0.0625	-0.1250
28	-0.4219	0.4063	-0.0625	0.2813	0.3437	-0.1875	-0.1250
29	0.0156	0.2187	0.1875	0.1562	-0.1563	-0.1250	-0.1875

COLUMN

ROW	22	23	24	25	26	27	28
1	-0.0938	0.1875	-0.0625	-0.0625	0.0937	0.0937	-0.2187
2	0.2188	0.1250	0.1250	0.0625	-0.0938	0.0937	0.0938
3	-0.2812	-0.3750	-0.0625	-0.1250	-0.0312	-0.0937	0.0937
4	0.1562	0.1875	0.0625	0.0625	0.0938	-0.0312	-0.0938
5	0.0937	0.1875	-0.0000	0.1250	-0.0313	0.0937	-0.0313
6	-0.2812	-0.1250	-0.0000	-0.0625	0.0313	-0.0312	0.0937
7	-0.0312	-0.3125	0.0625	-0.0625	0.0313	0.0313	-0.0313
8	0.1562	0.1875	-0.1250	0.1250	0.0312	0.0312	0.0313
9	0.0313	0.0625	0.0000	-0.1250	0.0312	0.0312	0.0313
10	0.0313	0.1250	0.0000	0.0625	-0.1563	0.0312	0.0312
11	0.0312	0.0000	0.0625	0.0000	0.0313	-0.1562	0.0313
12	0.0937	0.1250	-0.0625	0.1250	-0.0313	0.0312	-0.1563
13	0.0312	0.1250	-0.0625	-0.0000	0.0312	-0.0313	0.0313
14	-0.1562	-0.3125	-0.0625	-0.1875	-0.0937	0.0313	-0.0312
15	-0.0938	0.0000	-0.0000	-0.0625	-0.0312	-0.0938	0.0313
16	-0.0937	-0.1875	0.0625	-0.0625	-0.0312	-0.0312	-0.0937
17	-0.0312	0.0000	-0.0000	0.0625	0.0312	-0.0313	-0.0312
18	-0.0312	-0.2500	0.0625	-0.1250	0.0938	0.0313	-0.0313
19	0.1562	0.0625	-0.0625	0.0625	-0.0313	0.0938	0.0312
20	0.0937	0.0625	0.1250	-0.1250	0.0938	-0.0312	0.0938
21	-0.0312	0.2500	0.1250	0.1875	-0.0938	0.0938	-0.0313
22	-0.0313	0.0625	0.1875	0.1875	0.1562	-0.0937	0.0937
23	-0.0937	-0.1875	0.0000	0.1250	0.1563	0.1563	-0.0938
24	0.0312	0.1875	-0.1250	0.1250	0.1562	0.1562	0.1562
25	0.0313	-0.2500	-0.1250	-0.1875	-0.0312	0.1563	0.1562
26	0.0312	0.3750	-0.0625	0.0000	-0.0938	-0.0313	0.1562
27	0.0937	0.2500	0.0625	0.1250	-0.1563	-0.0938	-0.0313
28	0.2812	0.2500	-0.1875	0.2500	-0.0938	-0.1563	-0.0938
29	0.0937	0.0625	-0.1875	-0.1875	0.0313	-0.0938	-0.1562

COLUMN

ROW	29
1	0.0937
2	-0.3437
3	0.3437
4	-0.0312
5	-0.1562
6	0.0312
7	0.1563
8	-0.0937
9	0.0312
10	-0.0312
11	0.0313
12	-0.0313
13	-0.2187
14	0.1562
15	-0.0313
16	0.0937
17	-0.0938
18	0.0937
19	-0.0313
20	0.0937
21	0.0312
22	-0.0937
23	0.1562
24	-0.2187
25	0.2188
26	0.0313
27	-0.0312
28	-0.2187
29	-0.0937

References

1. Sloane, N. J. A., Fine, T., Phillips, P. G., and Harwit, M. (1969) Codes of Multiplex Spectrometry, *App. Opt.* **8**: 2103–2106.

2. Harwit, M., Phillips, P. G., Fine, T., and Sloane, N. J. A., Doubly Multiplexed Dispersive Spectrometers, *Appl. Opt.* (to be published).

3. Baumert, L. D. (1964) in *Digital Communications with Space Applications*, S. W. Golomb, Ed., Prentice-Hall, Englewood Cliffs, N.J., pp. 165–172.

Contents

48. A New Multiplexing Spectrometer with Large Throughput

Perry G. Phillips and Martin Harwit
Center for Radiophysics and Space Research
Cornell University
Ithaca, New York
and
Niel J. A. Sloane
Bell Telephone Laboratories, Inc.
Murray Hill, New Jersey

Abstract

A dispersion instrument which multiplexes the radiation at both its entrance and exit apertures is described. This double multiplexing scheme allows one to recover both Fellgett's advantage and the high throughput advantage normally attributed only to interferometric transform spectrometers. This instrument compares favorably with Michelson interferometric spectrometers. The first results of a laboratory pilot model are presented.

48–1 INTRODUCTION

Two types of multiplexing schemes that can be used for obtaining spectra have been known for about 20 years. The first of these was a scheme suggested by Golay[1] in which radiation at different frequencies is modulated at different rates by mechanical means. Such a system involves the dispersion of radiation by means of a grating or prism and the modulation of the dispersed radiation by a mechanical chopper or mask. The theory of a number of simple systems of this type has been further discussed by Ibbett *et al.*[2] Decker and Harwit,[3] and Sloane *et al.*[4] This type of spectrometer has, so far, not come into general use.

The second kind of multiplex spectrometer is the interferometric spectrometer, as exemplified by the Michelson interferometer discussed in detail by Fellgett[5] and by Vanasse and Sakai.[6] In recent years the interferometric spectrometers, particularly the Michelson spectrometer, have been very successful. One reason for this success lies in the very high energy throughput of the interferometric instruments. This advantage of the interferometers over grating instruments was clearly pointed out first by Jacquinot.[7] He showed that the throughput of the Fabry-Perot interferometer was far greater than that of a conventional grating instrument, (which, in turn, was greater than that of a prism spectrometer).

In defense of grating spectrometers, one should point out that an instrument of the type originally built by Golay did demonstrate high energy throughput and that perhaps Jacquinot's argument should be modified. As far as we know, this point has never been explicitly raised, presumably because the counter-argument can be raised that the multiplexing process introduced by Golay can also be used to improve interferometric spectrometers. For the grating spectrometer, multiplexing would thus increase luminosity, while for interferometers, such as those of the Fabry-Perot variety which already have high throughput, the corresponding advantage would lie in an ability to simultaneously observe all spectral wavelengths. The multiplexing process then converts a monochromatic device such as the Fabry-Perot

spectrometer into a polychromatic instrument like the Michelson. This multiplexing advantage is the so-called Fellgett's advantage.

The tacit assumption made in the above argument is that multiplexing can be used to achieve only one out of two possible advantages. Either one can have Golay's advantage in increased throughput, as is realized in instruments built by Golay, Girard[8] and others; or one can have Fellgett's advantage of polychromatic transmission. Either process increases the signal falling on the detector at any given time.

The point that perhaps has not been explicitly made before is that a double-multiplexing scheme can endow an instrument with both increased luminosity and polychromatic transmission.

The purpose of the present paper is to show that multiplex spectrometry with grating instruments actually permits the construction of high luminosity instruments that can compete with Michelson interferometric spectrometers. As with Michelson interferometers, these instruments show their greatest advantage under conditions where measurements are detector-noise limited and where the source of radiation is diffuse. With photon-noise limited detectors, only the luminosity gain is realized, but such a gain is still useful.

In the instrument analyzed here, an initial stage of multiplexing is introduced at the entrance of the spectrometer and a second multiplexing stage follows at the instrument's exit plane. The signal-to-noise ratio advantage, which can be achieved over an ordinary grating instrument, is shown to be comparable to that of the Michelson instrument. This can be understood as follows. In order to describe the intensity of N spectral elements, a Michelson interferometric spectrometer will generally need to make measurements at $\sim 2N$ different mirror separations (N measurements on each side of the white light fringes). Under certain conditions this number can be reduced to N. In a similar way, Sloane (Chapter 47) has shown that there are $(2N-1)$ determinations required for a doubly multiplexed grating instrument. The radiant power incident on the detector is roughly the same in both cases; it differs by a factor of the order of the ratio of the beamsplitter energy loss to the grating energy loss. This factor is on the order of unity and by no means amounts to the orders of magnitude claimed in comparisons of grating and interferometric instruments. Thus, there appears to exist no fundamental restriction which limits grating instruments to an inferior role, provided the multiplexing techniques are fully exploited.

48-2 MULTIPLEX ADVANTAGE

The encoding procedure has been described by Harwit *et al.*[9] and in Chapter 47 by Sloane. The former have shown that N^2 measurements can be made to estimate $2N-1$ spectral elements. The latter has indicated how the number of measurements can be reduced to $2N-1$. Table 48-1 gives the results of a comparison of the total mean square error for three different grating spectrometers. The first column gives the number of unknown spectral elements. The remaining three columns give the total mean square error for a conventional spectrometer, a singly multiplexed spectrometer described by Decker (Chapter 46), and a doubly multiplexed spectrometer making N^2 measurements described by Harwit *et al.*[9] The last two instruments use the S code (Chapters 46 and 47).

Table 48-1. Comparison of Total Mean Square Error for 3 Grating Spectrometers in Estimating N Unknowns

N	No Mask	Single Mask	Double Masks
3	$3\sigma^2$	$2.25\sigma^2$	$2.56\sigma^2$
7	$7\sigma^2$	$3.06\sigma^2$	$2.00\sigma^2$
11	$11\sigma^2$	$3.36\sigma^2$	$1.51\sigma^2$
19	$19\sigma^2$	$3.61\sigma^2$	$.98\sigma^2$
N	$N\sigma^2$	$\left(2-\frac{2}{N+1}\right)^2\sigma^2$	$\sim\frac{22.2}{N}\sigma^2$ for N large

(σ^2 is the mean square noise in a single measurement made in time $\frac{T}{N}$; $\sigma^2=\frac{\text{constant}\cdot N}{T}$.)

The comparison assumes: (1) a fixed total measuring time T; (2) constant incident energy density; (3) any equal number N of spectral elements to be estimated; (4) identical photodetectors; and (5) constant slot widths.

The only reason for presenting Table 48-1 on the basis of N^2 measurements, rather than $2N-1$, lies in the greater simplicity of the mathematical treatment of the error analysis. Note, however, that in astronomical applications the use of N^2 measurements can give additional information about the difference in the spectral distribution of light reaching the instrument from different portions of the source.

48-3 BENCH MODEL OF DOUBLY MULTIPLEXED SPECTROMETER

A bench model of a doubly multiplexed spectrometer was set up in the laboratory at Cornell. It is shown in Figure 48-1. Figure 48-2 gives the optical diagram. The masks were drawn to actual size and reproduced on clear plastic. A seven element S code was used with successive encoding positions generated in the stepwise manner described by Decker (Chapter 46). The two masks were fastened to two movable elements of a dove-tailed slide. Each slot of the mask was 1 mm wide giving a total slit width of 7 mm.

The collimating mirror was a 50 cm focal length, f/7 parabaloid. The dispersive element was a 7.5 cm square, 300 grooves/mm grating blazed at 1000 Å.

The system was aligned by imaging the entrance aperture on the exit aperture for the third order mercury green line. Other lines were excluded by a filter. The resolving power of the system was 925

Figure 48-1. Bench Model of Doubly Multiplexed Dispersive Spectrometer

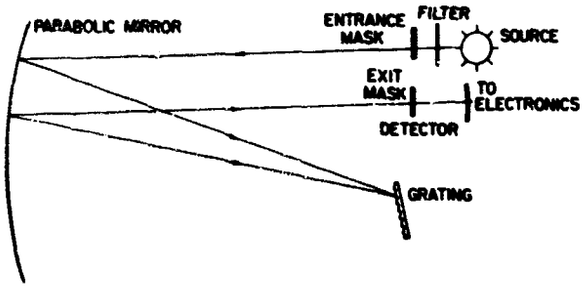


Figure 48-2. Optical Diagram of Bench Model Spectrometer

with an overall f-ratio of f/7. A phototube measured the transmitted radiation and a digital voltmeter displayed the phototube output.

The results of the experiment are shown in Figure 48-3. The dotted line is the recovered spectrum using N^2 ($N=7$) measurements employing the technique described by Harwit *et al.*,[9] and Sloane. The solid line is the spectrum recovered by using the linearly independent 13 measurements given by Sloane (Chapter 47).

The signal-to-noise ratio of the solid spectrum is worse because the total observing time was a factor of seven less than that of the broken spectrum. Nevertheless, the single mercury line is clearly visible in both cases.

Admittedly, the bench model was rather inelegant. Nevertheless, it points to the important fact that a doubly dispersive spectrometer is not plagued by the fine tolerances associated with interferometric spectrometers. It took only a few hours to set up the experiment and these results were produced the first time.

48-4 COMPARISON OF DOUBLY MULTIPLEXED AND MICHELSON SPECTROMETERS

The Michelson spectrometer was first compared to a single slit grating instrument. As is well known

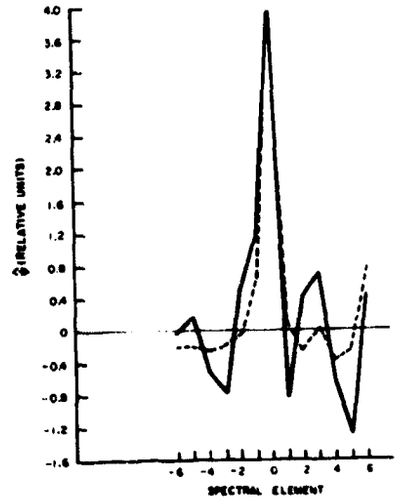


Figure 48-3. Spectra Recovered from Laboratory Bench Model Showing Mercury Green Line. The solid line is for $2N-1=13$ measurements and the broken line for $N^2=49$ measurements

the Michelson's advantages can be separated into two distinct kinds. First, there is a multiplex advantage, which, for N spectral elements, is equivalent to a reduction of the total mean square noise by a factor $\sim 2/N$. Second, there is the luminosity advantage which gives a greater effective signal for a constant detector noise. As viewed here, this will improve the figure of merit for the Michelson over that of a single slit grating instrument by another factor of the order of the ratio of the acceptance area times the acceptance solid angle of both instruments. For instruments with identical resolving powers, this ratio is not substantially different from the ratio of the throughputs of the doubly multiplexed and single slit instruments. A comparison[9] shows that when realistic grating and mask transmission losses are taken into account, a multiplexed grating instrument should still compare favorably with an interferometer in total energy throughput.

In other words, the Michelson interferometer's multiplex advantage gives an improvement in signal-to-noise ratio by a factor of order $N^{\frac{1}{2}}$. We find in Table 48-1 a similar improvement in going from a normal single slit spectrometer to a simple multiplexed system. The Michelson also has an energy throughput advantage over a single slit spectrometer. The additional gain of a factor of order N in the double mask situation of Table 48-1 gives a throughput advantage which also corresponds to that of the Michelson. A detailed comparison of the relative merits of interferometers and doubly multiplexed dispersive systems would therefore center around factors which would depend on such things as reflection and transmission losses, sizes of available gratings or beam splitters, and similar instrumental questions.

For these reasons, we feel that doubly multiplexed dispersive spectrometers should be seriously considered for high throughput applications.

References

1. Golay, M. J. E. (1949) *J. Opt. Soc. Am.* **39**: 437.

2. Ibbett, R. N., Aspinall, D., and Grainger, J. F. (1968) *Appl. Opt.* **7**: 1089.

3. Decker, J. A. Jr., and Harwit, M. O. (1968) *Appl. Opt.* **7**: 2205.

4. Sloane, N. J. A., Fine, T., Phillips, P. G. and Harwit, M. O. (1969) *Appl. Opt.* **8**: 2103.

5. Fellgett, P. (1958) *J. Phys.* **19**: 187.

6. Vanasse, G. A., and Sakai, H. (1967) *Progress in Optics* VI: 298.

7. Jacquinot, P. (1959) *J. Opt. Soc. Am.* **44**: 761.

8. Girard, A. (1963) *Appl. Opt.* **2**: 79 as quoted by J. T. Houghton and S. D. Smith (1966) in *Infra-red Physics*, Oxford at the Clarendon Press, p. 225.

9. Harwit, M., Phillips, P. G., Fine, T., and Sloane, N. J. A. (1970) *Appl. Opt.* to be published.

Contents

49. Degrees of Freedom and Computation Requirements in Matrix Multiplication for Hadamard and Other Spectrometry Transforms*

Harry C. Andrews
Department of Electrical Engineering
University of Southern California
Los Angeles, California 90007

Abstract

An algorithm is presented which enables certain matrix multiplications in a digital computer to be implemented with a considerable savings in storage and computational operations. For an $N \times N$ matrix vector multiplication a maximum of $N\sum_{r=0}^{n-1} P_r$ storage words and computer operations are necessary compared to normal matrix requirements of N^2 locations and operations. The algorithm has as subsets the generation of the fast Fourier transform, fast Hadamard transform, fast Walsh transform, fast Kronecker matrix transform, and an infinite class of transformations unnamed but potentially useful in generalised spectral analysis as well as coding, bandwidth reduction, and feature selection.

49–1 INTRODUCTION

The use of high speed digital computers has greatly enhanced the techniques of signal processing by allowing complex mathematical relations to be implemented in a relatively short span of time. However, certain computational tasks, such as matrix multiplication, still require an inordinate amount of computational complexity as well as storage demands. Because vector-matrix operations are so prevalent in signal analysis problems, it becomes quite useful to attempt a streamlining, where possible, of the computer matrix multiplication process. Probably the most famous of these efforts has resulted in the fast Fourier transform (FFT) in which the data-vector-Fourier-matrix multiplication is implemented much more effectively than normal matrix multiplication requirements.[1] Of somewhat lesser fame but of great importance is the fast Hadamard or Walsh transform (FHT, FWT) which also has resulted in considerable savings in computational and storage requirements for computer implementation.[2]

The techniques providing for the FFT and FHT are based upon a matrix factorization which appears to be due to Good[3] and which can be generalised to describe a factorization technique which is a function of the degrees of freedom in the definition of the particular matrix subject to multiplication. Thus, if an arbitrary $N \times N$ matrix is presented, it has a maximum of N^2 degrees of freedom and consequently

* This work was supported by a NASA Grant NGL-05-018-044 Supplement No. 3 and a JPL grant JPL-952334.

will require N^2 computer operations for vector-matrix multiplication. However, if the $N \times N$ matrix has some additional constraint limiting its maximum number of degrees of freedom, then it seems plausible that a vector matrix multiplication should require less than N^2 operations. Indeed, this paper presents a class of matrices of dimension $N \times N$ where N is a highly composite number,

$$N = \prod_{r=0}^{n-1} P_r,$$

P_r is an integer; and vector matrix multiplication can be implemented in

$$N \sum_{r=0}^{n-1} P_r$$

operations. The observant reader will note that when the set of $\{P_r\}$ are all the same the above dimension becomes $N = p^n$ and the number of operations becomes $pN \log_p N$ which is the common reference index for the FFT and FHT. However, in addition to having the FFT and FHT as subsets of the formulation presented here, the class of Kronecker matrices[4-6] are also included.

49-2 DEGREES OF FREEDOM

The development of an efficient matrix multiplication algorithm is based upon the amount of redundancy built into the matrix in question, and a measure of such redundancy is the maximum number of degrees of freedom allowable in the generation of the matrix. A degree of freedom is defined to be an arbitrary entry or parameter in a matrix. For the class of matrices to be described in this paper and denoted as a generalized fast transform (GFT) in keeping with the FFT and FHT, the maximum number of degrees of freedom is given by

$$N \sum_{r=0}^{n-1} P_r$$

and N has the aforementioned constraint of

$$N = \prod_{r=0}^{n-1} P_r.$$

The GFT matrix formulation (also described as a GFT machine) is designed to implement a class of matrices described by a maximum of

$$N \sum_{r=0}^{n-1} P_r$$

degrees of freedom in the same number of computational operations. Thus, such a machine is optimal in some sense since it would be impossible to implement a linear transformation in fewer operations than degrees of freedom. In addition to implementating the multiplication in fewer number of computations than normally required (N^2), it is possible to save considerable storage requirements in the computer. This storage savings results from storing only the parameters necessary for algorithmic generation of the matrix and will always be equal to the actual degrees of freedom and considerably less than N^2 storage requirements for our arbitrary $N \times N$ matrix. It will be necessary to develop algorithms for the description of the matrices in arithmetic rather than two dimensional array formats. Towards this end it becomes useful to investigate lexicographic number systems.

49-3 LEXICOGRAPHIC NOTATION

Lexicographic or dictionary notation refers to the use of a variety of bases or radices in a positional digit notation system. Thus, the numbers represented by X and U can be described as

$$X = X_{n-1}X_{n-2}, \ldots, X_1X_0$$

$$\text{where } X_r \in \{0, 1, \ldots, P_r - 1\}, \quad (49\text{-}1)$$

and

$$U = U_{n-1}U_{n-2}, \ldots, U_1U_0$$

$$\text{where } U_r \in \{0, 1, \ldots, P_r - 1\}. \quad (49\text{-}2)$$

The base or radix is represented by the integers P_r and the rth digit counts modulo

$$\prod_{k=0}^{r-1} P_k.$$

The rth digit will always cycle through P_r unique integers from 0 through $P_r - 1$.

In order that advantage can be made of the lexicographic notation, it becomes useful to denote the rows of matrices by the index X, and the columns of matrices by the index U. Then, to describe a particular element within a matrix, it is only necessary to note the value of the row X and column U intersection. Note that both X and U take on values zero through $N-1$. It becomes necessary to decode the indices in a digit-by-digit fashion similar to binary bit decoding. Towards this end consider a mathematical technique of decoding the X variable defined as

$$X = X_{n-1}, \ldots, X_r, \ldots, X_0 \; X_r \in \{0, 1, \ldots, P_r - 1\}, \tag{49-3}$$

to be the operation of

$$\prod_{r=0}^{n-1} \sum_{i=0}^{P-1} \delta(X_r - i). \tag{49-4}$$

Here the delta function, $\delta(a-b)$, takes on the value unity only when $a=b$ and is zero otherwise. The digit-by-digit decoding operation characterized by the mathematical product and sum relations above can be described as summing all the delta functions possible for the rth digit, only one of which will be nonzero. The product operand then multiplies the correct delta function for each digit to describe the entire word X. Therefore, the decoding expression will be unity only when the X value has been decoded properly and will be zero otherwise.

49-4 MATRIX FORMULATION

The underlying premise for implementing fast algorithms and efficient storage techniques for multiplication of redundantly defined matrices is to take advantage of the maximum number of degrees of freedom of submatrices generating the ultimate $N \times N$ matrix. Thus, if a total of n core matrices M_r are defined, such that each core matrix is of dimension $P_r \times N$, then the total maximum number of degrees of freedom for a set of n core matrices becomes

$$N \sum_{r=0}^{n} P_r.$$

This implies that the GFT will be completely defined by n core matrices M_r, each of dimension P_r by N and it will then be necessary to store only

$$N \sum_{r=0}^{n-1} P_r$$

parameters rather than the N^2 entries of the transformation matrix.

While advantage has been taken of the reduced storage requirement, generation of the resulting transformation has yet to be demonstrated. The ultimate transformation matrix, denoted by $H_n(X, U)$, will be of dimension $N \times N$ and implies that the entries at the Xth row and Uth column will be algorithmically defined. In addition, the $H_n(X, U)$ matrix will be factorable into a product of n matrices denoted by G_r, each of which will have only $P_r N$ nonzero entries and will be completely defined by the respective core matrix M_r. Therefore

$$H_n(X, U) = [G_{n-1}][G_{n-2}] \ldots [G_r] \ldots [G_1][G_0]. \tag{49-5}$$

The fact that the transformation matrix $H_n(X, U)$ is factorable into a product of matrices G_r, each of dimension $N \times N$ and each with a high density of zero elements, implies that a reduced number of operations will be necessary for vector matrix multiplication of the series of nG_r matrices as opposed to the $H_n(X, U)$ matrix of possibly nonzero entries. The nonzero entries in the G_r matrices are those in the M_r core matrices and are deterministically placed such that there are only P_r nonzero entries in each row and in each column. Therefore when a vector is multiplied by the first G_{n-1} matrix, a total of $P_{n-1}N$ operations will be necessary. When the resulting vector is multiplied by the next matrix, a total $P_{n-2}N$ operations will be necessary. After all vector-G_r-matrix operations are complete, a total of

$$N \sum_{r=0}^{n-1} P_r$$

operations will have been involved, a considerable savings over the normal vector $H_n(X, U)$ matrix multiplication requirement of N^2 operations. Each G_r matrix, denoted as a Good matrix after the author

of one of the original factorization techniques,[3] is completely described by the respective M_r core matrix.

The ultimate transformation matrix $H_n(X, U)$ is given by the matrix product in Eq. (49-5). However, this equation can be simplified according to the decoding notation introduced earlier:

$$z=\left(\sum_{y_r=0}^{P_r-1}\delta(X_r-y_r)\right)\left(\prod_{s=0}^{r-1}\sum_{v_s=0}^{P_s-1}\delta(X_s-v_s)\right)\left(\prod_{s=r}^{n-1}\sum_{v_s=0}^{P_s-1}\delta(U_s-v_s)\right) \qquad (49\text{-}6)$$

and

$$H_n(X, U)=\prod_{r=0}^{n-1} m_r^{z}{}_{,y_r,v_r} \qquad (49\text{-}7)$$

Equations (49-6) and (49-7) can be interpreted in the following way. The entry in a given row X and column U is determined by the product of n parameters. Each of the n parameters originates from a different core matrix. In all cases the column index v of the core matrix is completely defined, n digits, as

$$G_2=\begin{bmatrix}1&&&&1&&&\\&1&&&&1&&\\&&1&&&&1&\\&&&1&&&&1\\1&&&&-1&&&\\&1&&&&-1&&\\&&1&&&&-1&\\&&&1&&&&-1\end{bmatrix}\cdot G_1=\begin{bmatrix}1&&1&&&&&\\&1&&1&&&&\\1&&-1&&&&&\\&1&&-1&&&&\\&&&&1&&1&\\&&&&&1&&1\\&&&&1&&-1&\\&&&&&1&&-1\end{bmatrix}\cdot G_0=\begin{bmatrix}1&1&&&&&&\\1&-1&&&&&&\\&&1&1&&&&\\&&1&-1&&&&\\&&&&1&1&&\\&&&&1&-1&&\\&&&&&&1&1\\&&&&&&1&-1\end{bmatrix}$$

$$H_n=[G_2]\,[G_1]\,[G_0]$$

$$H_n=\begin{bmatrix}1&1&1&1&1&1&1&1\\1&-1&1&-1&1&-1&1&-1\\1&1&-1&-1&1&1&-1&-1\\1&-1&-1&1&1&-1&-1&1\\1&1&1&1&-1&-1&-1&-1\\1&-1&1&-1&-1&1&-1&1\\1&1&-1&-1&-1&-1&1&1\\1&-1&-1&1&-1&1&1&-1\end{bmatrix}$$

Figure 49-1. 8×8 Hadamard Matrix

$$M_2=\begin{bmatrix}W^0&W^0&W^0&W^0&W^0&W^0&W^0&W^0\\W^0&W^0&W^0&W^0&W^4&W^5&W^6&W^7\end{bmatrix},\quad M_1=\begin{bmatrix}W^0&W^0&W^0&W^2&W^0&W^0&W^0&W^2\\W^0&W^0&W^4&W^6&W^0&W^0&W^4&W^6\end{bmatrix},\quad M_0=\begin{bmatrix}W^0&W^0&W^0&W^0&W^0&W^0&W^0&W^0\\W^0&W^4&W^0&W^4&W^0&W^4&W^0&W^4\end{bmatrix}$$

$$G_2=\begin{bmatrix}W^0&-&-&-&W^0&-&-&-\\-&W^0&-&-&-&W^1&-&-\\-&-&W^0&-&-&-&W^2&-\\-&-&-&W^0&-&-&-&W^3\\W^0&-&-&-&W^4&-&-&-\\-&W^0&-&-&-&W^5&-&-\\-&-&W^0&-&-&-&W^6&-\\-&-&-&W^0&-&-&-&W^7\end{bmatrix},\quad G_1=\begin{bmatrix}W^0&-&W^0&-&-&-&-&-\\-&W^0&-&W^2&-&-&-&-\\W^0&-&W^4&-&-&-&-&-\\-&W^0&-&W^6&-&-&-&-\\-&-&-&-&W^0&-&W^0&-\\-&-&-&-&-&W^0&-&W^2\\-&-&-&-&W^0&-&W^4&-\\-&-&-&-&-&W^0&-&W^6\end{bmatrix},\quad G_0=\begin{bmatrix}W^0&W^0&-&-&-&-&-&-\\W^0&W^4&-&-&-&-&-&-\\-&-&W^0&W^0&-&-&-&-\\-&-&W^0&W^4&-&-&-&-\\-&-&-&-&W^0&W^0&-&-\\-&-&-&-&W^0&W^4&-&-\\-&-&-&-&-&-&W^0&W^0\\-&-&-&-&-&-&W^0&W^4\end{bmatrix}$$

$$H_2=[G_2][G_1][G_0]$$

$$H_2=\begin{bmatrix}W^0&W^0&W^0&W^0&W^0&W^0&W^0&W^0\\W^0&W^4&W^2&W^6&W^1&W^5&W^3&W^7\\W^0&W^0&W^4&W^4&W^2&W^2&W^6&W^6\\W^0&W^4&W^6&W^2&W^3&W^7&W^1&W^5\\W^0&W^0&W^0&W^0&W^4&W^4&W^4&W^4\\W^0&W^4&W^2&W^6&W^5&W^1&W^7&W^3\\W^0&W^0&W^4&W^4&W^6&W^6&W^2&W^2\\W^0&W^4&W^6&W^2&W^7&W^3&W^5&W^1\end{bmatrix},\quad \overline{H}_2=\begin{bmatrix}W^0&W^0&W^0&W^0&W^0&W^0&W^0&W^0\\W^0&W^1&W^2&W^3&W^4&W^5&W^6&W^7\\W^0&W^2&W^4&W^6&W^0&W^2&W^4&W^6\\W^0&W^3&W^6&W^1&W^4&W^7&W^2&W^5\\W^0&W^4&W^0&W^4&W^0&W^4&W^0&W^4\\W^0&W^5&W^2&W^7&W^4&W^1&W^6&W^3\\W^0&W^6&W^4&W^2&W^0&W^6&W^4&W^2\\W^0&W^7&W^6&W^5&W^4&W^3&W^2&W^1\end{bmatrix}$$

Lexicographically Reordered

$$\overline{H}_2=\left[\exp\frac{2\pi iux}{8}\right]$$

$$W=\exp\left(\frac{2\pi i}{8}\right)$$

Figure 49-2. 8×8 Fourier Transform

is the row index y_r of the respective core matrix. Note that not all digits of the row X or column U indices are defined by the decoding in the exponent. This simply means that the particular core entry defined by the decoding occurs a multiple number of times in the $H_n(X, U)$ matrix in the locations not completely defined by the missing digits in the X and U lexicographic representations. This, indeed, is what allows each entry of the M_r core matrix to affect N/P_r different entries in $H_n(X, U)$.

While Eqs. (49-6) and (49-7) are mathematically cumbersome, it is important to realize that they only represent decoding of lexicographically represented row and column indices. Consequently, the algorithmic generation of the $H_n(X, U)$ transformation matrix in a computer program will only decode the lexicographic row and column indices digit by digit.

49–5 APPLICATIONS

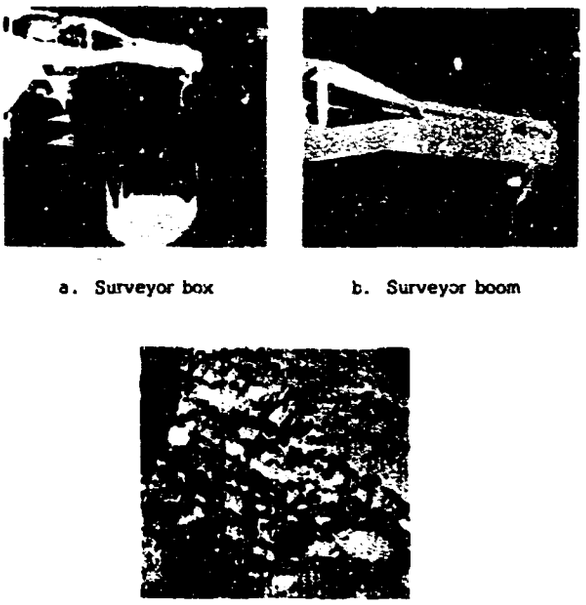
It has been shown that the formulation presented above includes the class of transformations defined by Kronecker products.[6] Probably the simplest example of a transformation resulting from a Kronecker product is that given by a 2×2 matrix Kroneckered with itself n times. Let

$$M_r = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{49-8}$$

The 8×8 matrix generated by this core is in Figure 49-1. This matrix is known as a Hadamard matrix[7] of order 2 and implements a discrete Walsh transform on a vector multiplied by the matrix.[8]

As a final example, consider the generation of a Fourier transform of resolution 8×8. In this case three core matrices of 2×8 dimensions each will completely define the transformation. Figure 49-2 presents such an example. Notice the need for lexicographic reordering to recombine the spectral coefficients according to increasing frequency. This need manifests itself in the FFT literature[1] and the reordering in this example is nothing more than reading the binary representation of the columns of H_2 in reverse order resulting in the $\overline{H}_2$ matrix.

Examples of two-dimensional transformations using both the Fourier and Hadamard GFT implementation are presented with the original scenes in Figures 49-3, 49-4, and 49-5.



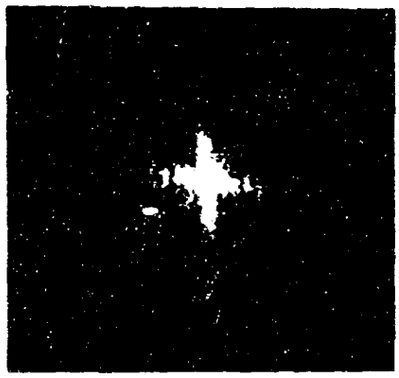
a. Surveyor box b. Surveyor boom

c. Moonscape

Figure 49-3. Original Images

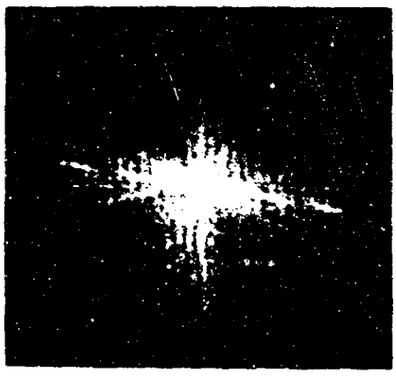
49–6 CONCLUSIONS

This paper presents the results of research into a method of streamlining certain vector matrix multiplication operations for implementation in a digital computer. The efficient storage and implementation requirements have been obtained based on a degrees-of-freedom analysis and it has been shown that the algorithm for multiplication of matrices described in this paper is optimum in the sense that there are as many possible degrees of freedom as computer operations necessary (there could be no more degrees of freedom for the same number of computer operations). The algorithm presented is based on a lexicographic notation system such that digit-by-digit decoding enables one to keep track of how a few numbers of parameters control a large number of entries in the ultimate transformation matrix. The algorithm for description of the transformation matrix is presented in a rather complex equation which is shown to considerably simplify for restrictive cases. Such cases included as subsets of this algorithm are the class of Kronecker matrices, fast Fourier transforms (FFT), and fast Walsh or Hadamard transforms (FWT, FHT), and a variety of other matrix transformations. Specific matrix and image transform examples are presented to verify the analysis.



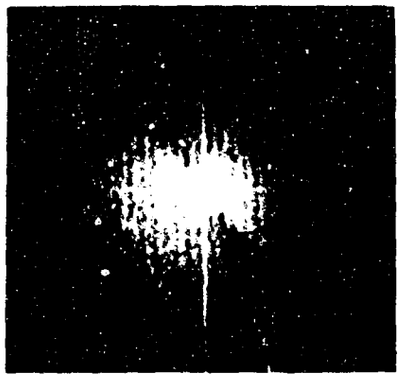
a. Logarithm display of Surveyor box transform

b. Threshold display of Surveyor box transform

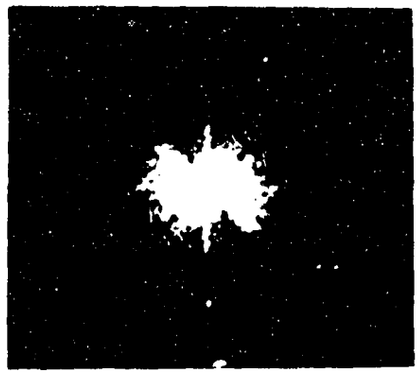


c. Logarithm display of Surveyor boom transform

d. Threshold display of Surveyor boom transform

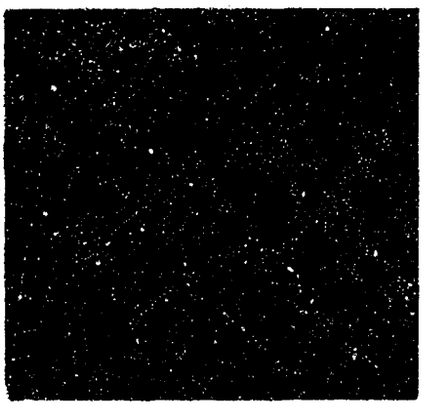


e. Logarithm display of Moonscape transform

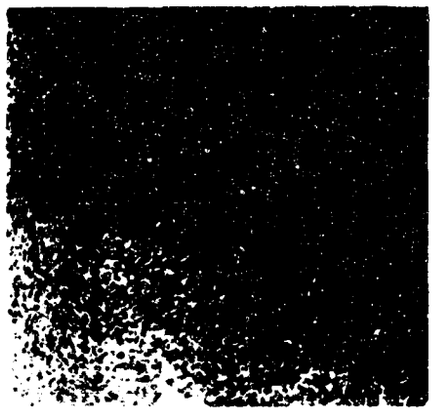


f. Threshold display of Moonscape transform

Figure 49-4. Fourier Transforms



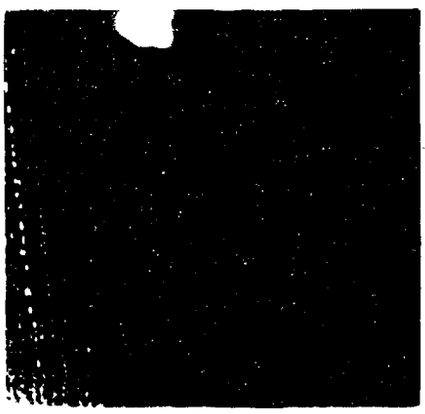
a. Logarithm display of Surveyor box transform



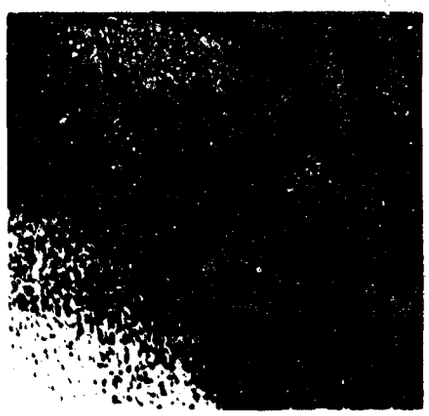
b. Threshold display of Surveyor box transform

c. Logarithm display of Surveyor boom transform

d. Threshold display of Surveyor boom transform



e. Logarithm display of Moonscape transform



f. Threshold display of Moonscape transform

Figure 49-5. Hadamard Transforms

References

1. Cooley, J. W., and Tukey, J. W. (1965) An Algorithm for the machine calculation of complex Fourier series, *Math. Computation* **19**: 297–301.

2. Whelchel, J. E., and Guinn, D. F. (1968) The fast Fourier-Hadamard transform and its use in signal representation and classification, *Eascon '68 Record*, pp. 561–573.

3. Good, I. J., (1958) The interaction algorithm and practical Fourier analysis, *J. Roy. Stat. Soc.* **820**: 361.

4. Bellman, R. (1960) *Introduction to Matrix Analysis*, McGraw-Hill, New York, p. 227.

5. Andrews, H. C., and Kane, J. (1970) Kronecker Products, Computer Implementation, and Generalized Spectra, *JACM*.

6. Andrews, H. C., and Caspari, K. L. (1970) A generalized technique for spectral analysis, *IEEE Trans. on Computers*.

7. Pratt, W. K., Kane, J., and Andrews, H. C. (1969) Hadamard transform image coding, *Proc. IEEE* **57** (No. 1): 58–68.

8. Walsh, J. L. (1923) A closed set of normal orthogonal functions, *An. J. Math.* **45**: 5–24.

Contents

50. Fourier Spectroscopy at the Lunar and Planetary Laboratory of the University of Arizona

Uwe Fink and Harold Larson
University of Arizona
Tucson, Arizona

Abstract

The high altitude spectroscopy program of the Lunar and Planetary Laboratory at the University of Arizona is briefly outlined. A short description of the instrument and the data reduction technique is given. Representative spectra of the Moon, Venus and Mars, taken aboard the high altitude airplane, are shown. Further spectra taken with a ground-based telescope are then shown and discussed.

50–1 INTRODUCTION

The Fourier spectroscopy program started at the Lunar and Planetary Laboratory when a small 20 cm^{-1} resolution interferometer was made available to Dr. G. P. Kuiper by L. Mertz.[1,2] The interferometer was used on one of the flights, in April 1967, of the NASA Convair 990 airplane as part of the Lunar and Planetary Laboratory high altitude spectroscopy program. Despite aircraft vibrations the instrument performed very well. Spectra were obtained of Mars and Venus with the Moon as a comparison during seven flights in the period April to June 1967. These early spectra together with a description of the high altitude program are published in the *Communications of the Lunar and Planetary Laboratory*.[3,4,5]

The purpose of high altitude flights is to get above the tropopause, the cold trap for water vapor in our atmosphere. The tropopause lies near altitude 7 to 8 km in the polar areas, 11 to 12 km in the middle latitudes, and 18 km in the tropics. Thus, for effective use in our latitudes, a plane must fly higher than 12 km or about 40,000 ft. The water vapor then remaining above the plane is remarkably independent of season or local weather conditions, being about 10μ precipitable H_2O in the direction of the zenith.[6] Thus, the water vapor is cut by more than a factor of 1000 for near sea level operation, or by almost a factor of 100 for a good infrared mountain observatory site. The effect of the decreasing water vapor is shown quite dramatically in the solar spectra in Figure 50-1 taken on 23 June, 1967, using the 20 cm^{-1} resolution Mertz interferometer. Some absorption remains in the 1.9 μ H_2O band at 41,000 ft altitude, a little less for the 1.4 μ band, while the absorption at 1.1 μ is for all practical purposes eliminated.

From the Venus spectra obtained during the same series of flights, it was immediately possible to put a lower or upper limit for water vapor or ice clouds in the Venus atmosphere.

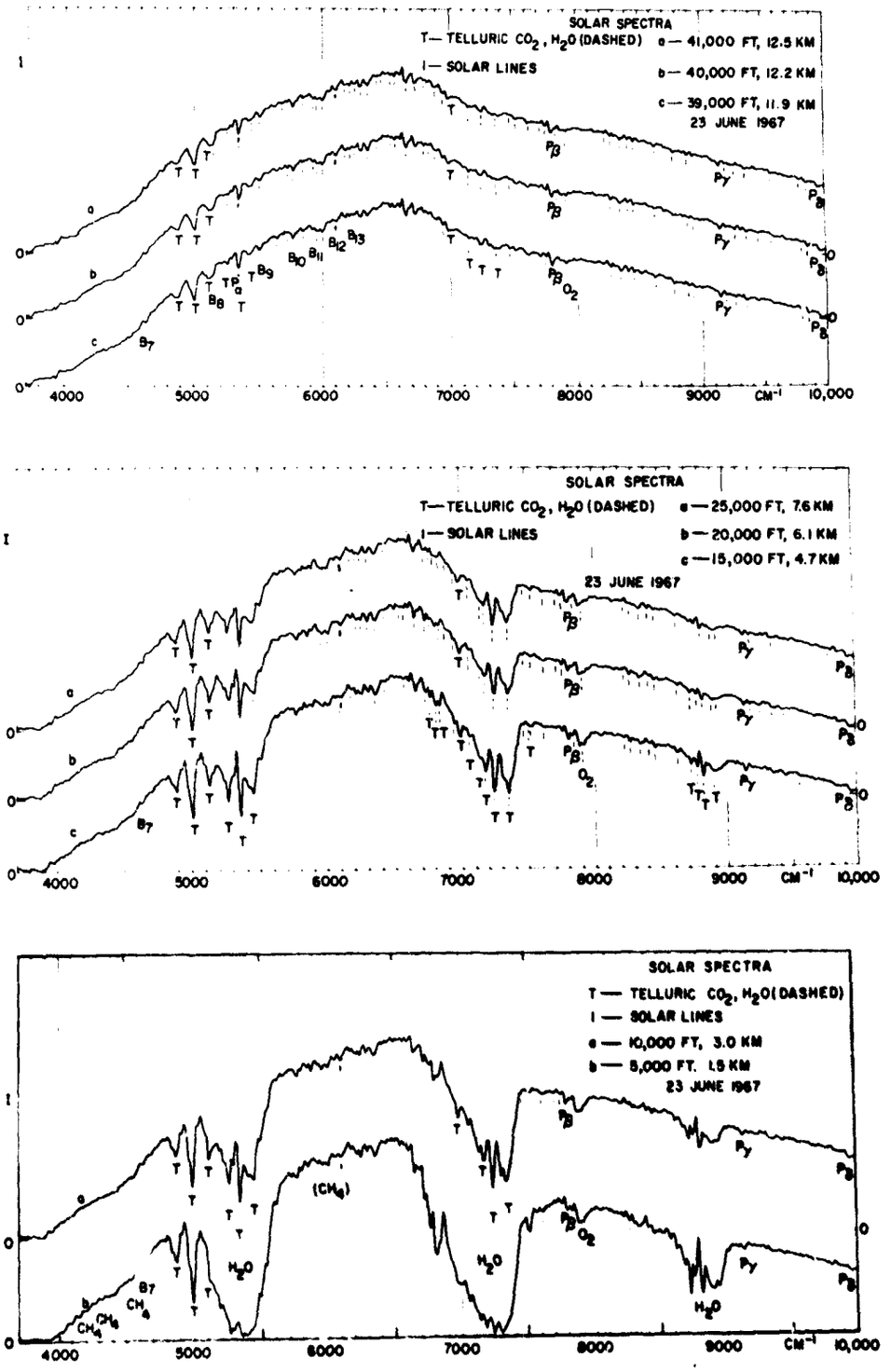


Figure 50-1. Solar Spectra Taken on Board the NASA CV 990 Jet Based at Moffett Field with a 20 cm^{-1} Resolution Interferometer. The spectra show the decrease in water vapor absorption in the 1.1 μ, 1.4 μ, and 1.9 μ bands as the plane climbs above the tropopause

50-2 INTERFEROMETER DESCRIPTION AND DATA REDUCTION SYSTEM

Due to the success of the 20 cm^{-1} interferometer, a new interferometer of 8 cm^{-1} resolution was obtained from Block Associates. The interferometer is of the rapid scan type with an f/13 input beam to match existing telescopes. The instrument, together with the PbS detector cold box, is shown in Figure 50-2, attached at the Cassegrain focus of the 61-in.

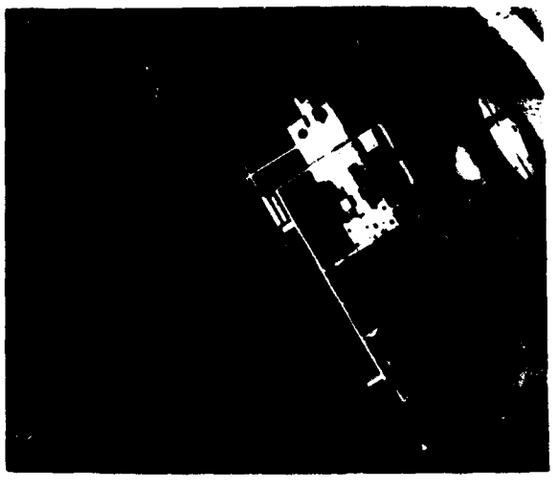


Figure 50-2. The Interferometer Attached to the Cassegrain Focus of the 61-in. Catalina Observatory Telescope (Dr. Kuiper observing)

Catalina Observatory telescope of the Lunar and Planetary Laboratory. In the airplane it was used with a 12-in. telescope fed by a gyrostabilized heliostat mirror looking through a 12×14 in. dry quartz window at an angle of 65° from the horizon.

The interferometer consists of two "cubes", that is, essentially two separate interferometers with the moving mirrors in both cubes driven by the same magnetic coil. The "reference cube" has the He 1.09 μ line as a monochromatic input and a white light input. The "signal cube" has the potential for two inputs and two outputs. The center of the signal interferogram is adjusted so that it occurs about 120 μ after that of the white light interferogram of the reference cube.

The instrument was modified by Dr. H. L. Johnson in several ways. The mirror displacement was slightly increased to give a resolution of about 5 cm^{-1}. A broadband green input was added to the signal cube. With this additional interferogram occurring at the same place as the signal interferogram, small temperature dependent phase shifts between the two cubes could be monitored and corrected. The two uncooled PbS detectors were replaced by a single, good 0.25×0.25 mm PbS detector cooled to liquid N_2 temperature. The beam was focused on the detector by a system of a CaF_2-LiF achromatic triplet and a sapphire aplanatic immersion lens designed by Dave Steinmetz.

The velocity of the moving mirror is such that the frequency of the 1.08 μ reference line is about 800 Hz. The output from the preamp and amplifier of the detector goes through a filter set at 150 to 900 Hz (corresponding to the spectral region of about 1.5 μ). It is then recorded along with the reference, white and green signals, on a seven-track analog magnetic tape recorder.

The system of adding the interferograms was also changed from the original Block coadder and paper tape to a procedure developed jointly by Roger Thomson of MIT and R. I. Mitchell of the University of Arizona. The magnetic tape is read at twice the recording speed through another filter and an AD converter into an IBM 1130 computer. The computer adds the interferograms in phase by means of a trigger from the signal cube white light. The zero crossings of the reference fringes provide the sample points for the AD converter. The relative drift between the two cubes is adjusted by means of the green signal. (The reference cube and signal cube signals are read with two different heads of the tape recorder so that a change of the tape distance between the heads with a micrometer is equivalent to a phase change between the two cubes.) In the case of the airplane data, about 1000 to 2000 scans were added in each case to give a coadded interferogram. A phase corrected power spectrum is obtained from the Fourier transform of the interferogram.

50-3 HIGH ALTITUDE AND GROUND-BASED SPECTRA

Samples of the spectra obtained with the 8 cm^{-1} resolution interferometer during the high altitude program are shown in Figures 50-3, 50-4, and 50-5. Figure 50-3 shows a spectrum of the Moon obtained during two flights on May 5 and 6, 1968, with a total observing time of five hours. The remains of the telluric water vapor bands at 1.4 μ, 1.9 μ, and 2.7 μ are indicated by dots above the lines. Most of the other lines remaining are solar lines, the strongest of which are identified on the figure. Special note should be made of the Bracket lines of hydrogen B6 to B13.

Figure 50-4 is a ratio of the spectrum of Venus taken on November 27 and 28, 1967, divided by the spectrum of the Moon of Figure 50-3. Total observing time for Venus was 3.3 hours. The many CO_2 bands, which are of course shown so much more clearly in the Connes' spectra[7] can be seen in this figure. The spectrum of the Moon of Figure 50-3 was used for the ratio since it was a very clean spectrum and would not introduce additional noise in the ratio spectrum. Moon spectra were also taken on the Venus flights with very nearly the same air mass as Venus. From these spectra an abundance of 2.5 μ precipitable H_2O or a mixing ratio with respect to CO_2 of 10^{-6} was obtained. A full description of this analysis is given in Kuiper *et al.*[8] In Table 50-1 is presented a summary

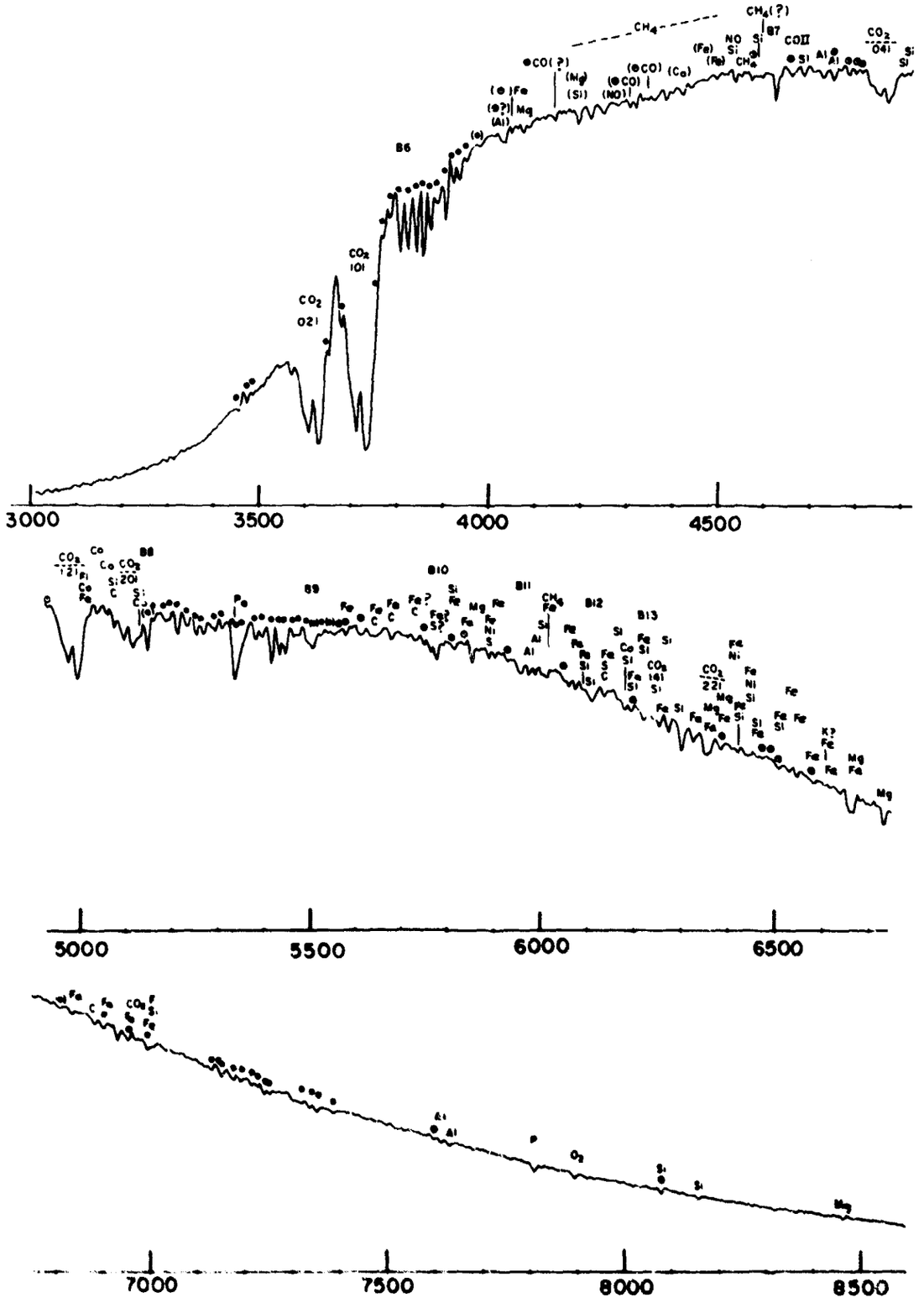


Figure 50-3. Lunar Comparison Spectrum, CV 990 May 5 and 6, 1968 at About 41,000 ft Altitude. A number of prominent solar lines are identified. The remaining water vapor lines corresponding to about 10 μ precipitable H_2O are indicated by dots

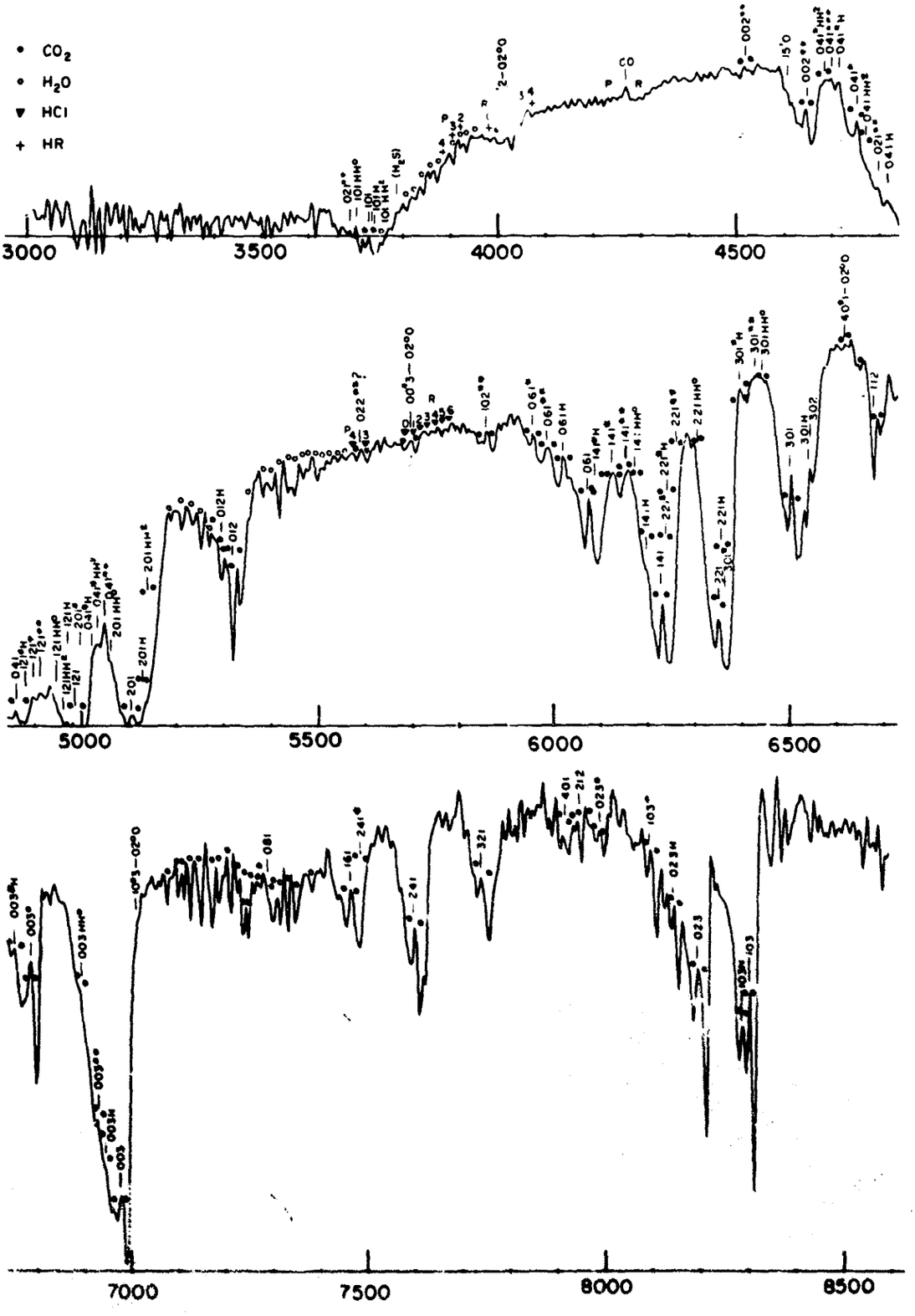


Figure 50-4. The Airplane Ratio Spectrum of Venus for Nov. 27 and 28, 1967, Divided by the Spectrum of the Moon for May 5, 6, 1968. The identification of the large number of CO_2 bands is given. Water vapor lines are designated by open circles

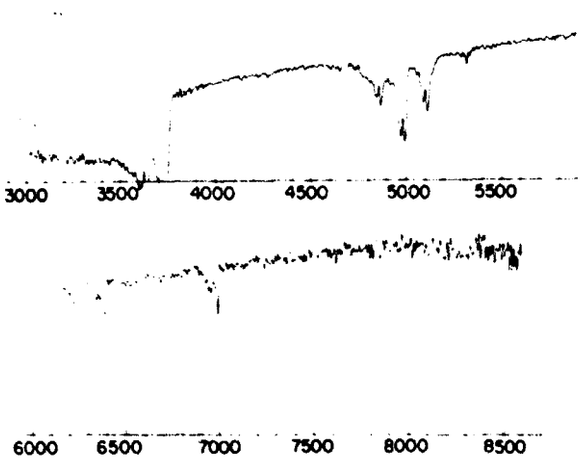


Figure 50-5. The CV 990 Airplane Ratio Spectrum of Mars for May 29, 30, June 2, 4 Divided by the Spectrum of the Moon for May 26, 29, 30. Water vapor is marginally detectable on Mars in the H_2O band at 3800 cm^{-1}

of the present knowledge of the gaseous constituents of the Venus atmosphere. The table reflects favorably on Fourier interferometry since all of the entries by Connes and many of the Lunar and Planetary Laboratory entries were obtained using this method, together with laboratory calibrations.

An airplane ratio of the spectrum of Mars taken May 29, 30, and June 2, 4, 1969, divided by the spectrum of the Moon for May 26, 29, 30, 1969, is shown in Figure 50-5. The Mars flights comprise 8 to 9 hours of observing time. The CO_2 bands on this figure are, of course, much weaker than on Venus but considerably stronger than the telluric ones in the Moon spectra. Unfortunately, even averaging four flights does not quite provide sufficient signal-to-noise ratio to make possible a good measurement of the Martian water vapor. Only in the strong, saturated 2.7 μ band are there any water lines left in the ratio.

Ground based spectra were taken with the Catalina Observatory 61-in. telescope. Figure 50-6 shows

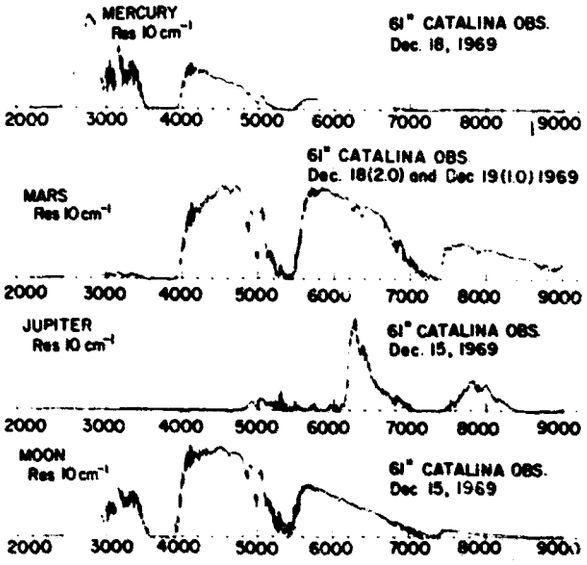


Figure 50-6. The Spectra of Mercury, Mars, Jupiter, and the Moon Taken With the 61-in. Catalina Observatory Telescope. The fine structure seen in the spectra is real and requires a larger scale for effective display

Table 50-1. Composition of Venus Atmosphere

Gas	Mixing Ratio	Source
CO_2	0.95 ± 0.02	Venera 4, 5, 6 (Pravda, June 4, 1969)
N_2	0.035 ± 0.015	Venera 4, 5, 6 (Pravda, June 4, 1969)
CO	$10^{-4.34}$	Connes *et al.*, *Ap. J.*, 152, p. 741, 1968
HCl	$10^{-6.2}$	Connes *et al.*, *Ap. J.*, 147, pp. 1230–1237, 1967
HF	$10^{-8.2}$	Connes *et al.*, *Ap. J.*, 147, pp. 1230–1237, 1967
CH_4	$< 10^{-6}$	Connes *et al.*, *Ap. J.*, 147, p. 1235, 1967
CH_3Cl	$< 10^{-6}$	Connes *et al.*, *Ap. J.*, 147, p. 1235, 1967
CH_3F	$< 10^{-6}$	Connes *et al.*, *Ap. J.*, 147, p. 1235, 1967
C_2H_2	$< 10^{-6}$	Connes *et al.*, *Ap. J.*, 147, p. 1235, 1967
HCN	$< 10^{-6}$	Connes *et al.*, *Ap. J.*, 147, p. 1235, 1967
H_2O	$10^{-6.0}$	Kuiper *et al.*, *Comm. LPL* No. 100
O_2	$< 10^{-5.0}$	Belton and Hunten, *Ap. J.*, 153, p. 970, 1968
O_3	$< 10^{-8}$	Jenkins *et al.*, *Contrib. KPNO* No. 421, 1969
SO_2	$< 10^{-7.5}$	Cruikshank and Kuiper, *Comm. LPL* No. 97
COS	$< 10^{-6}$	Cruikshank, *Comm. LPL* No. 98
CS	$< 10^{-8}$	Based on *Comm. LPL* No. 100
C_3O_2	$< 10^{-6.3}$	Based on *Comm. LPL* No. 100
H_2S	$< 10^{-3.7}$	Cruikshank, *Comm. LPL* No. 98
H_2S		Based on *Comm. LPL* No. 100
NH_3	$< 10^{-7.5}$	Based on *Comm. LPL* No. 100

spectra of Mercury, Mars, Jupiter, and the Moon. The full resolution of all these interferograms is 5 cm^{-1}, but they are shown transformed only to 10 cm^{-1} resolution in the figure. They serve to orient the reader to what types of spectral features are to be expected for these planets. The spectra of the Moon show the typical strengths of the telluric H_2O absorption bands that can be expected at the 8260 ft (2.50 km) altitude of the Catalina Observatory. Mars shows the well known CO_2 atmosphere. Jupiter is almost opaque in the infrared due to CH_4 and NH_3 absorptions. No atmosphere on Mercury was detected. The previous upper limit of 0.04 mb of CO_2 on Mercury[9] can probably be lowered by means of our spectra.

Figure 50-7 shows a spectrum of Saturn and the

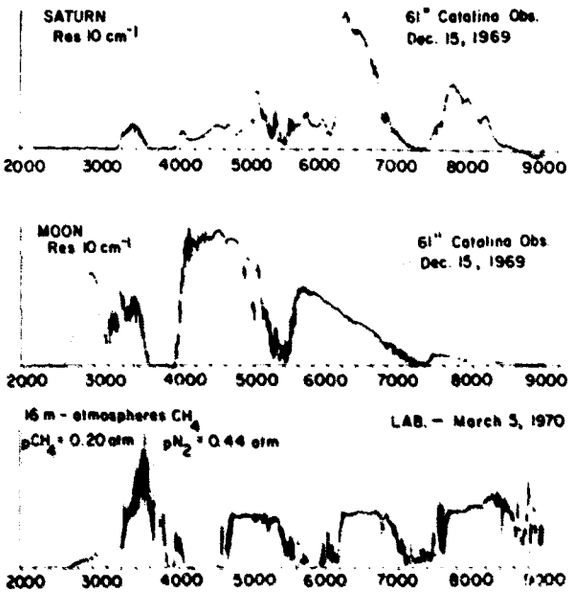


Figure 50-7. The Spectra of Saturn and the Moon Compared With Laboratory Spectrum of CH_4 With the 40 M White Cell. The match is not complete due to the effect of temperature on the intensity of the bands and the scattering atmosphere of Saturn

Moon. An attempt to interpret the Saturn spectrum was made by taking laboratory comparison spectra with the 40 m white cell. A sample is shown on the bottom of Figure 50-7. A faint channel spectrum originating in the interferometer can also be detected. The match with Saturn is not quite perfect due to the temperature difference and the scattering properties of Saturn's atmosphere. However, no constituent other than CH_4 was detected with the possible exception of the feature at 4800 cm^{-1} which may be due to the S(1) line of the pressure induced dipole spectrum of molecular hydrogen.[10]

A spectrum of the rings of Saturn, shown in Figure 50-8, together with a spectrum of the Moon

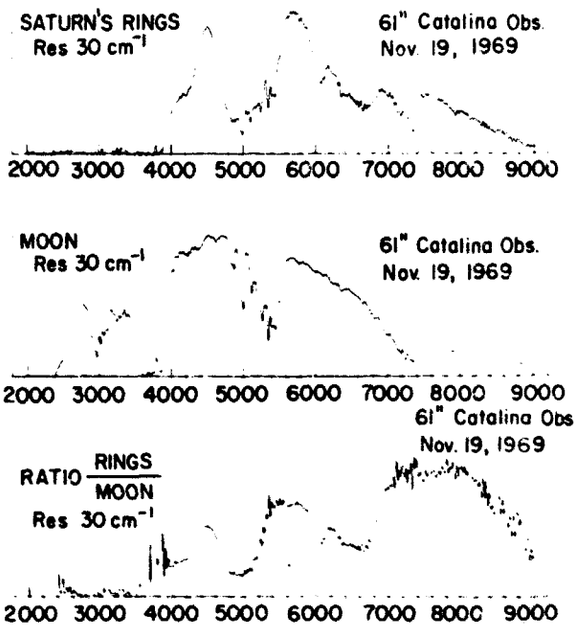


Figure 50-8. The Spectra of Saturn's Rings, the Moon, and the Ratio Spectrum

obtained on the same night and reduced at the same resolution was also obtained. Their ratio is shown below. The night was a particularly dry one as can be seen from the small absorption of the 1.4 μ and 1.8 μ water vapor bands. In a search for the interpretation of the spectrum frozen NH_3 (Figure 50-9) which showed a sharp feature at 6100 cm^{-1} agreeing with the ring spectrum, was first considered. Ordinary ice at approximately $-20°C$ did not seem to fit.[11] It was then pointed out by a group at MIT[12] that water ice might present a better fit. Further spectra were taken in the laboratory and the temperature dependence of H_2O frost indicated in Figure 50-10 and previously unsuspected was found. The ice at liquid N_2 temperature, shown on a slightly larger scale in Figure 50-10, provides a good match as is indicated by the ratio at the bottom of the figure. Further work on this problem is in progress.

In addition to the planetary spectra just discussed, a number of stellar spectra were obtained by Dr. H. L. Johnson and his coworkers with the interferometer. These were taken with the 5-ft metal mirror telescope of the Lunar and Planetary Laboratory, Catalina Observatory. Spectra of carbon stars, M type stars, K type stars, Sirius, and α Tauri were obtained.[13,14,15] They led to the identification of SiO by Jean Fertel.[16]

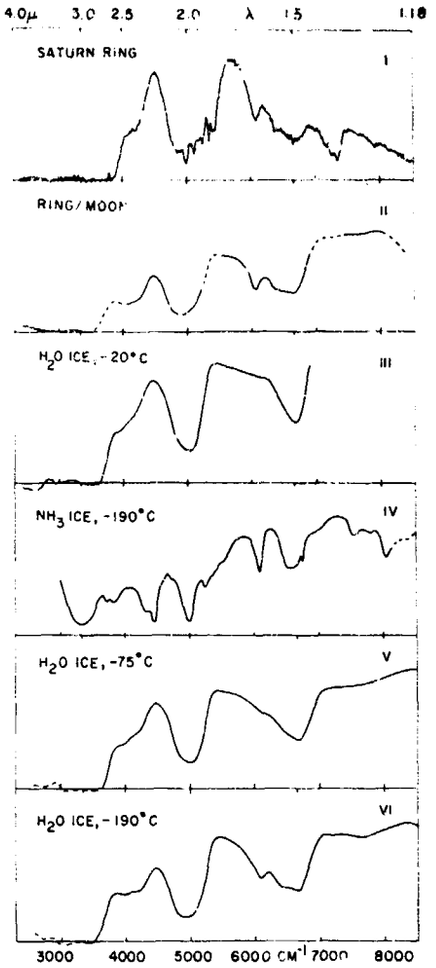


Figure 50-9. Comparison of the Spectrum of the Rings of Saturn to Various Laboratory Ices

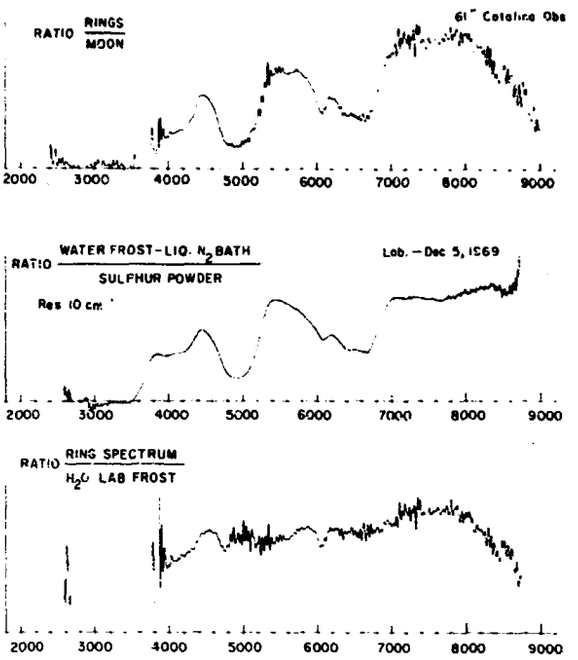


Figure 50-10. Comparison of the Absorption Features Seen in Saturn's Rings to Those of Water Ice at Liquid N_2 Temperatures

50–4 CONCLUSIONS

In the Fourier spectroscopy program at the Lunar and Planetary Laboratory concentration was on two aspects of this new technique: (1) The multiplex advantage has been used on faint objects at low resolution where the weak signal does not permit higher resolution; (2) The interferometer has been used successfully in environmental conditions somewhat more severe than in the laboratory, as for example the high altitude jet observation platform.

In this program the small size and weight of the instrument, which permits it to be readily attached to the Cassegrain focus of a moderate size telescope, has been helpful. The success of the airplane observations may, to a large extent, be due to the rapid scan technique.

Acknowledgments

The program of planetary spectroscopy at the Lunar and Planetary Laboratory is supported in part by NASA under contract NsG 161-61.

References

1. Mertz, L. (1965) *Astron. J.* **70**: 548.

2. Mertz, L. (1965) *Transformations in Optics*, John Wiley and Sons, New York.

3. Kuiper, G. P., Forbes, F. F., and Johnson, H. L. (1967) *Commun. Lunar and Planetary Lab.* **6** (No. 93): 155.

4. Kuiper, G. P., and Steinmetz, D. L. (1967) *Commun. Lunar and Planetary Lab.* **6** (No. 94): 171.

5. Kuiper, G. P., and Forbes, F. F. (1967) *Commun. Lunar and Planetary Lab.* **6** (No. 95): 177.

6. Kuhn, P. M. ESSA, private communication, Boulder, Colorado.

7. Connes, J., Connes, P., and Maillard, S. P. (1969) *Near Infrared Spectra of Venus, Mars, Jupiter and Saturn*, CNRS, Paris.

8. Kuiper, G. P., Forbes, F. F., Steinmetz, D. L., and Mitchell, R. I. (1969) *Commun. Lunar and Planetary Lab.* **6** (No. 100): 209.

9. Bergstralk, J. T., Gray, L. P., and Smith, H. J. (1967) *Ap. J.* **149**: 137.

10. Welsh, H. L. (1969) *J. Atm. Sci.* **26**: 835.

11. Kuiper, G. P., Cruikshank, D. P., and Fink, Uwe (1970) *Sky and Telescope* **39**: 14.

12. Pilcher, C. B., Chapman, C. R., Lebofsky, L. A., and Kieffer, H. N. (1970) *Science* **167**: 1372.

13. Thomson, R. I., Schnopper, H. W., Mitchell, R. I., and Johnson, H. L. (1969) *Ap. J. Letters* **158**: L55.

14. Thomson, R. I., Schnopper, H. W., Mitchell, R. I., and Johnson, H. L. (1969) *Ap. J. Letters* **158**: L117.

15. Johnson, H. L., Coleman, I., Mitchell, R. I., and Steinmetz, D. L. (1968) *Commun. Lunar and Planetary Lab.* **7** (No. 113): 83.

16. Fertel, J. H. (1970) *Ap. J. Letters* **158**: L7.

Discussion

Q. (Robert Dick, Barringer Research, Ltd.): Does Jupiter's spatial structure have any effects on your measurements?

A. (Uwe Fink): The spectra that were shown were taken with the whole planet so that we could not resolve any structure. We are, however, planning to take spatially resolved spectra of the bands on Jupiter.

Q. (Robert Dick): You use the moon as reference. Are you confident that the moon's surface does not introduce any special features?

A. (Uwe Fink): The albedo of the moon, especially if one includes the emission at large wavelengths, is not completely flat. It will not, however, introduce any sharp spectral features of the type we are looking for. It is necessary to have a comparison for the telluric water vapor within a few hours of the planet, and the moon is most convenient for that.

Q. (D. A. Otmen, University of Alberta): Have you considered the possibility that the spectra of Saturn's rings might be caused by clathrates of ice? These are entities in which molecules of one substance are completely enclosed in a lattice of another, a water lattice in this case.

A. (Uwe Fink): No, we haven't.

Q. (D. A. Otmen): These substances have spectra which are similar to those of ice, and their spectra are currently being investigated and seem to show extra lines due to the guest molecules. This might allow identification of other molecules in Saturn's rings.

A. (Uwe Fink): The spectral features that we see on the ring spectrum are all very well explained by low temperature ice, and we do not, with the present resolution, see any extra lines.

Q. (Dr. Alain L. Fymat, Jet Propulsion Laboratory): Do you plan to record spectra inside and outside of the Jupiter's red spot for abundance determinations? This would represent a partial test to Hide's Taylor column explanation of the spot.

A. (Uwe Fink): Yes, we plan to do that at the same time that we scan the belts of Jupiter.

Q. (Henry U. Blau, Jr., Arthur D. Little, Inc.): How did you make the low temperature frosts?

A. (Uwe Fink): The frosts were made in a cylindrical section of glass taking about 5 cm in. diam and 10 cm long. Both ends were closed with glass plates and the cell was pumped out. It was then immersed in a dewar of liquid N_2 and water vapor was admitted by boiling distilled water from a flask at a pressure of a few μ of Hg.

Aspen ICFS '70 Exhibits

255-SLOT HADAMARD-TRANSFORM SPECTROMETER

Comstock & Wescott, Inc., 765 Concord Avenue, Cambridge, Massachusetts 02138 and AFCRL: John A. Decker, Jr., organizer.

Demonstration model of a 255-slot Hadamard-Transform Spectrometer operating on the principle of multislit optical encoding at spectrometer exit focal plane. The encoded optical data are punched on data tapes and then fed into a PDP-10 time sharing computer for decoding and automatic plotting of the output spectra. The instrument was operated at the Conference.

FOURIER SPECTRA RECORDED AT LABORATOIRE AIMÉ COTTON

Centre National Recherche Scientifique, 91 Orsay, France: P. Connes, organizer.

Exhibit of Fourier spectra recorded at Laboratoire Aimé Cotton and calculated at CIRCE, including the originals of figures for three papers presented at the Conference. This means 10^6 samples emission (thorium, holmium) or absorption (N_2O, KH_3) spectra, selected portions of astronomical spectra (Venus and Mars, with the lines of CO, HCl, and HF), and a copy of the Planetary Atlas. The first results given by the C.N.R.S. real-time computer were also presented.

AIRBORNE INTERFEROMETERS FOR MEASURING INFRARED ATMOSPHERIC EMISSION SPECTRA

Air Force Cambridge Research Laboratories and Stewart Radiance Laboratory, Utah State Univ., Bedford, Massachusetts: A. T. Stair, Jr., E. Ray Huppi, and Ronald J. Huppi, organizers.

An interferometer used for measuring infrared atmospheric radiation was shown. The interferometer is one of three units that take measurements in the 1 to 15 μ wavelength region from an Air Force KC-135 aircraft. The three units are mounted behind a sophisticated cold chopper system that makes it possible to obtain sky-emission measurements far below the background levels of the ambient aircraft optics.

Some spectral measurements of the night sky were shown. The data were taken at various altitudes and temperatures. As a result, the variations in thermal emission of the atmosphere were shown.

The vibration isolator mounts and temperature control circuits required to make the interferometers flyable were also demonstrated.

AIRBORNE INTERFEROMETER SPECTROMETER

Air Force Cambridge Research Laboratories, Bedford Massachusetts: John D. Rex and John P. Cahill, organizers.

Airborne Interferometer Spectrometer used on AFCRL KC-135 aircraft.

Detector—InAs

Sweep Rate—2 interferograms/sec

Resolution—10 cm^{-1}

Recent eclipse data (coronal emission in 1.5 to 3.3 μ region, 7 March 70) were shown along with atmospheric transmission measurements in the 3 to 7.5 μ region.

GRUBB-PARSONS MH11 FOURIER SPECTROMETER

Edwin Industries Corporation, 11933 Tech Road, Silver Spring, Maryland 20904: D. Stewart Larson, organizer.

A Grubb-Parsons MH11 Fourier Spectrometer capable of producing absorption spectra in the range of 10 to 800 cm^{-1}, connected to a computer for real-time calculations of spectra from interferograms. The system will display any piece of the spectrum between 10 and 800 cm^{-1} and continually update it as the path difference in the interferometer is increased. Maximum resolution = 0.1 cm^{-1}.

MICHELSON AND LAMELLAR GRATING FOURIER SPECTROPHOTOMETERS

Beckman Instruments, 2500 Harbor Blvd., Fullerton, California 92634: G. T. Keahl, organizer.

Featured Michelson and Lamellar Grating Fourier Spectrophotometers and accessories along with analog and digital computer systems. The Michelson interferometer (developed by Beckman-R.I.I.C., London, England) is a modular system designed for use in the 10 to 500 cm^{-1} region with resolutions up to 0.1 cm^{-1}. The Lamellar Grating interferometer covers the range 3 to 70 cm^{-1} with resolutions up to 0.1 cm^{-1}. A new gas cell was exhibited with externally adjustable path lengths up to 10 m.

AN ENCODING MASK

Sterrewacht "Sonnenborgh" Servaasbolwerk 13, Utrecht, Holland: Thys M. W. M. de Graauw, organizer.

The mask is on a rotating cylinder. The spectrum is focused at the mask and the part of interest in the spectrum is divided into 255 elements. Every element is chopped by a pseudo-random function, and all the functions are shifted one clock pulse to each other. In this case, the maximum length sequence of 255 elements was used. Decoding means cross-correlation of the used sequence with the integrated signal. The cross-correlation function gives directly the intensity distribution in the observed spectrum.
The exhibited chopper was used for getting line-profiles of the photosphere-chromosphere transition region during the solar eclipse in 1970. The spectral elements had a spatial width of 80 $\mu m \pm \frac{1}{2}$ μm.

DEMONSTRATION FOURIERSCOPE

National Standards Laboratory, Sydney, N.S.W. 2008 Australia: W. H. Steel, organizer.

A direct-vision instrument that presents side-by-side the spectrum of a light source and the corresponding interferogram. These may be changed by placing optical filters in the instrument, so demonstrating qualitatively the properties of a Fourier transform.

PROTOTYPE IR15 INSTRUMENT FOR NIMBUS III

Texas Instruments, Inc., 13500 N. Central Expressway, P.O. Box 6015, MS 202, Dallas, Texas 75222: R. D. Hanel, NASA/GSFC, and D. H. Rodgers, Texas Instruments, organizers.

The exhibit consisted of the prototype of the IR15 instrument flown on NIMBUS III, together with its electronics plus a display of some of the results obtained from the NIMBUS flight. The significant features of this instrument are its thermal stability, mechanical integrity, and self-calibration features. The instrument will retain useful alignment over a 60°C temperature range without any adjustment. It will also withstand the qualification levels of vibration for the NIMBUS satellite, which puts forces up of to 40 g on the optical components. Finally, the instrument automatically calibrates itself every 16 scans by viewing deep space, followed by a view of an on-board target whose temperature is accurately known. Improved versions of this instrument are to be flown on the NIMBUS D and MARINER '71 spacecraft.

Attendees

K. D. Abhyankar
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109

Michael F. A'Hearn
University of Maryland
College Park, Maryland 20762

Francis J. Ahern, Jr.
University of Maryland
College Park, Maryland 20762

Ralph Alexander
University of Missouri
Physics Department
Rolla, Missouri 65401

Richard C. Altrock
Sacramento Peak Observatory
Sunspot, New Mexico 88349

Charles R. Anderson
Los Alamos Scientific Laboratory
Los Alamos, New Mexico 87544

Kenneth L. Andrew
Purdue University
Lafayette, Indiana 47909

Harry C. Andrews
University of Southern California
Los Angeles, California 90007

James R. Aronson
Arthur D. Little, Inc.
15 Acorn Park
Cambridge, Massachusetts 02140

Robert C. Arzbaecher
University of Illinois at Chicago
P.O. Box 4348
Chicago, Illinois 60680

Glen W. Ashley
General Dynamics
1675 W. Missouri
Pomona, California 41766

Gordon C. Augason
Ames Research Center
Moffett Field
Mountain View, California 94035

Wayne Bailey
University of Arizona
Steward Observatory
Tucson, Arizona 85721

Doran J. Baker
Utah State University
Electro-Dynamics Laboratories
Logan, Utah 84321

J. Ballantyne
School of Electrical Engineering
Cornell University
Ithaca, New York 14850

D. Boyd Barker
University of Denver
Denver, Colorado 80210

Harold Michael Barnes, Jr.
National Air Pollution Control Administration
5710 Wooster Pike
Cincinnati, Ohio 45227

Michael K. Barnoski
Hughes Research Laboratories
3011 Malibu Canyon Road
Malibu, California 90265

Joseph J. Barrett
Perkin-Elmer Corporation
50 Danbury Road
Wilton, Connecticut 06897

James P. Baukus
Cornell University
Electrical Engineering Department
Ithaca, New York 14850

Ely E. Bell
Ohio State University
Physics Department
Columbus, Ohio 43210

Robert J. Bell
University of Missouri
Physics Department
Rolla, Missouri 65401

David A. Benfield
University of Texas
Astronomy Department
Austin, Texas 78712

Yvon Biraud
Department Infra-Rouge Spatial
Observatoire de Paris-Meudon
92 Meudon
FRANCE

William E. Blass
University of Tennessee
Physics Department
Knoxville, Tennessee 37916

Henry U. Blau, Jr.
Arthur D. Little, Inc.
15 Acorn Park
Cambridge, Massachusetts 02140

Joe P. Blea
University of Missouri
Rolla, Missouri 65401

Hans J. Bolle
Meteorologisches Institut
11B Pflegerstr.
8 Munich 60 GERMANY

Murk Bottema
Ball Brothers Research Corporation
Boulder, Colorado 80401

Patrick Bouchareine
Laboratoire Aimé Cotton
C.N.R.S.
91 Orsay, FRANCE

Jack Brannain
Honeywell, Inc.
1625 Zarthan Avenue
St. Louis Park, Minnesota 55416

Serge Bras
Faculte des Sciences de Paris
9-Quai Saint-Bernard
Paris Veine, FRANCE

James W. Brault
Kitt Peak National Observatory
950 N. Cherry Avenue
Tucson, Arizona 85717

Anthony J. Broderick
Idealab, Inc.
Union St.
Franklin, Massachusetts 02038

Fon Brown, Jr.
Utah State University
Logan, Utah 84321

Marshall H. Bruce
AFCRL (OPR)
Bedford, Massachusetts 01730

Jay Burch
Naval Research Laboratory
4555 Overlook Dr. S.W.
Washington, D.C. 20390

Kenneth J. Button
M.I.T.
National Magnet Laboratory
170 Albany Street
Cambridge, Massachusetts 02139

John P. Cahill
AFCRL (OPR)
Bedford, Massachusetts 01730

Robert Calfe
ESSA Research Laboratory
Boulder, Colorado 80302

Nicholas P. Callas
Air Force Office of Scientific Research
1400 Wilson Blvd.
Arlington, Virginia 22209

Richard Canfield
Sacramento Peak Observatory
Sunspot, New Mexico 88349

Subhasn Chandra
General Electric
Space Sciences Laboratory
P.O. Box 8555
Philadelphia, Pennsylvania 19101

Robert J. Chambers
University of California
Astronomy Department
Berkeley, California 94720

Lucian W. Chaney
University of Michigan
Research Activities Bldg.
Hayward & Beal
Ann Arbor, Michigan 48106

Harry J. Chase
Lockheed Electronics Company
Manned Spacecraft Center
Houston, Texas 77571

Kenneth D. Chickering
Kollmorgen Corporation
104 Lower Westfield Road
Holyoke, Massachusetts 01040

Vern L. Chupp
Cary Instruments
2724 S. Peck Road
Monrovia, California 91016

Leroy L. Cogger
Arecibo Observatory
P.O. Box 995
Arecibo, Puerto Rico 00612

Isaiah Coleman
Block Engineering, Inc.
19 Blackstone Street
Cambridge, Massachusetts 02139

Janine Connes
C.N.R.S.
91 Orsay
FRANCE

Pierre Connes
C.N.R.S.
91 Orsay
FRANCE

Arthur Corman
Stewart Radiance Laboratory
40 North Road
Bedford, Massachusetts 01730

Dale P. Cruikshank
Lunar and Planetary Laboratory
University of Arizona
Tucson, Arizona 85721

Adam Csoeke-Poeckh
Denver Research Institute
University of Denver
Denver, Colorado 80903

Max Cuisenier
C.N.R.S.
Observatoire de Meudon
92 Meudon
FRANCE

Cameron Cumming
Defense Research Establishment
Valcartier P.O. Box 880
Courceletter, P.Q
CANADA

Raul Curbelo
Digilab, Inc.
237 Putnam Avenue
Cambridge, Massachusetts 02139

Malcom Currie
Beckman Instruments, Inc.
2500 Harbor Blvd.
Fullerton, California 92634

Don Davis
Time/Data Corporation
490 San Antonio Road
Palo Alto, California 94306

John Decker, Jr.
Comstock and Wescott, Inc.
765 Concord Avenue
Cambridge, Massachusetts 02138

Thys M. W. M. de Graauw
Sterrewacht "Sonnenborgh"
Servaasbolwerk 11–13
Utrecht, Holland

Luc D. DelBouille
Institut d'Astrophysique
B-4200 Cointe-Ougrie
BELGIQUE

Herve Delouis
C.N.R.S.
91 Orsay
FRANCE

Alvin M. Despain
Utah State University
Electro-Dynamics Laboratories
Logan, Utah 84321

C. Forbes Dewey, Jr.
M.I.T.
Room 3–258
Cambridge, Massachusetts 02139

Robert Dick
Barringer Research Ltd.
Rexdale, Ontario, CANADA

Jerome M. Dowling
Aerospace Corporation
Space Physics Laboratory
P.O. Box 95085
Los Angeles, California 90045

S. T. Dunn
Block Engineering, Inc.
11601 Newport Mill Street
Silver Spring, Maryland 20902

John A. Eddy
High Altitude Observatory
Box 1558
Boulder, Colorado 80302

T. Harvey Edwards
Michigan State University
Physics Department
East Lansing, Michigan 48823

Philip N. Eisner
The Dewey Electronics Corporation
311 East 47th St.
New York, New York 10017

Amos Englesrath
NASA—Ames Research
Moffett Field
Mountain View, California 94035

Wayne F. J. Evans
University of Saskatchewan ISAS
Saskatoon, Saskatchewan

Paul D. Feldman
Johns Hopkins University
Physics Department
Baltimore, Maryland 21218

Aaron S. Filler
623 Woodlawn St.
Fall River, Massachusetts 02720

Frank D. Findlay
Defence Research Establishment
Valcartier
P.O. Box 1427 Quebec
Quebec, CANADA

Uwe Fink
Lunar and Planetary Laboratory
University of Arizona
Tucson, Arizona 85721

Dale J. Flinders, Colonel, USAF
AFCRL (CC)
Bedford, Massachusetts 01730

Fred F. Forbes
University of Arizona
Optical Sciences Center
Tucson, Arizona 85721

Charles Foskett
Digilab, Inc.
237 Putnam Avenue
Cambridge, Massachusetts 02139

Alain L. Fymat
Jet Propulsion Laboratory
4800 Oak Grove Dr.
Pasadena, California 91103

Helene S. Gabelnick
Chicago State College
6800 So. Stewart
Chicago, Illinois 60621

Stephen D. Gabelnick
Argonne National Laboratory
Chemical Engineering Division
9700 S. Cars Avenue
Argonne, Illinois 60439

William A. Gale
Bellcomm, Inc.
955 L'Enfant Plaza, SW
Washington, D.C. 20024

James J. Gallagher
Martin Marietta Corporation
1005 Keats Avenue
Orlando, Florida 32809

William S. Gallawan
Beckman Instruments, Inc.
2500 Harbor Blvd.
Fullerton, California 92632

John S. Garing
AFCRL (OP)
Bedford, Massachusetts 01730

Hugh A. Gebbie
National Bureau of Standards
Boulder, Colorado 80302

William L. Gebel
State University of New York
Department of Earth & Space Sciences
Stony Brook, New York 11790

R. Ross Getty
Michigan State University
East Lansing, Michigan 48823

Richard E. Gillespie
Martin Marietta Corporation
P.O. Box 179
Denver, Colorado 80201

Andre Girard
O.N.E.R.A.
92 Chatillon
FRANCE

Lawrence Giver
NASA—Ames Research
Moffett Field
Mountain View, California 94035

Jacob H. Goldstein
Emory University
Chemistry Department
Atlanta, Georgia 30322

William D. Gregory
Georgetown University
Physics Department
Washington, D.C. 20007

Robert N. Grenda
General Electric Company
P.O. Box 8555
Philadelphia, Pennsylvania 19101

G. Guelachvili
Laboratoire Aimé Cotton
C.N.R.S.
91 Orsay
FRANCE

Richard T. Hall
Space Physics Laboratory
Aerospace Corporation
P.O. Box 95085
Los Angeles, California 90045

Rudolf A. Hanel
NASA Goddard Space Flight Center
Greenbelt, Maryland 20771

Niels-Peter R. Hansen
University of Massachusetts
Astronomy Research Facility
Amherst, Massachusetts 01002

Oswaldo M. Harris
Lunar and Planetary Observatory
University of Arizona
Tucson, Arizona 85721

Ronald Haswell
Grubb-Parsons
Newcastle-Upon-Tyne
NE62YB ENGLAND

Bernard F. Hawker
Beckman R.I.I.C. Ltd.
Worsley Bridge Road
London SE 26 ! LAND

Geoffrey Hayward
Beckman R.I.I.C. Ltd.
Worsley Bridge Rd.
London, ENGLAND

John C. Hill
General Motors Research Laboratory
18645 Wildemere
Detroit, Michigan 48221

Merle N. Hirsh
Dewey Electronics Corporation
311 E. 47th St.
New York, New York 10017

Bernard F. Hochheimer
Applied Physics Laboratory
8621 Georgia Avenue
Silver Spring, Maryland 20910

Joseph E. Hoffman, Jr.
465 Crown St.
Meriden, Conn. 06450

Marcus P. Hogue
Willey Optical Corporation
635 W. NASA Blvd.
Melbourne, Florida 32901

John N. Howard
AFCRL (CA)
Hanscom Field
Bedford, Massachusetts 01730

Susanne Hummer
Denver Research Institute
University of Denver
Denver, Colorado 80210

E. Ray Huppi
AFCRL (OPR)
Hanscom Field
Bedford, Massachusetts 01730

Ronald J. Huppi
Stewart Radiance Laboratory
40 North Road
Bedford, Massachusetts 01730

Robert B. Hutchinson
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, California 91103

Pierre Jacquinot
C.N.R.S.
Laboratoire Aimé Cotton
91 Orsay, FRANCE

Charles Y. Johnson
Naval Research Laboratory
Code 7127
Washington, D.C. 20390

James L. Jones
Applied Research Laboratories
Hasler Research Center
95 La Patera Lane
Goleta, California 93017

Gerald T. Keahl
Beckman Instruments, Inc.
2500 Harbor Blvd.
Fullerton, California 92632

John P. Kennealy
AFCRL (OPR)
Hanscom Field
Bedford, Massachusetts 01730

John Kline
Colorado School of Mines
Physics Department
Golden, Colorado 80401

Irving L. Kofsky
Photo Metrics, Inc.
442 Marrett Road
Lexington, Massachusetts 01273

Peter Krohmer
Bodenseewerk Perkin-Elmer & Co.
7770 Überlinger/Bodensee
Postfach 1120, WEST GERMANY

Ralph H. Kummler
Wayne State University
Detroit, Michigan 48202

Haydee K. Kurlat
University of Tennessee
Physics Department
Knoxville, Tennessee 37916

Miguel Kurlat
University of Tennessee
Physics Department
Knoxville, Tennessee 37916

Thomas G. Kyle
University of Denver
Denver, Colorado 80210

Russell D. Larsen
Illinois Institute of Technology
Chemistry Department
Chicago, Illinois 60616

James L. Lauer
Sun Oil Company
P.O. Box 426
Marcus Hook, Pennsylvania 19061

Robert H. Lee
High Altitude Observatory
P.O. Box 1558
Boulder, Colorado 80302

Lutz Leopold
Georgetown University
Physics Department
Washington, D.C. 20007

George Leroi
Chemistry Department
Michigan State University
East Lansing, Michigan 48823

Judah Levine
National Bureau of Standards
Boulder, Colorado 80302

Gilbert M. Levy
Barringer Research Ltd.
Rexdale, Ontario, CANADA

E. L. Lewis
Joint Institute for Laboratory Astrophysics
University of Colorado
Boulder, Colorado 80302

Gary W. Lindberg
Utah State University
Electro-Dynamics Laboratories
Logan, Utah 84321

Herbert G. Lipson
AFCRL (LQO)
Hanscom Field
Bedford, Massachusetts 01730

Robert L. Little
Beckman Instruments Inc.
7500 Harbor Blvd.
Fullerton, California 92634

Vincent J. Llamas
University of Missouri
Rolla, Missouri 65401

Ernest V. Loewenstein
AFCRL (OPI)
Hanscom Field
Bedford, Massachusetts 01730

Robert M. MacQueen
High Altitude Observatory
P.O. Box 1558
Boulder, Colorado 80302

John Pierre Maillard
C.N.R.S.
Laboratoire Aimé Cotton
91 Orsay, FRANCE

William G. Mankin
High Altitude Observatory
P.O. Box 1558
Boulder, Colorado 80302

Oscar P. Manley
Visidyne, Inc.
Woburn, Massachusetts 01801

Ham W. Marshall
Perkin-Elmer Corporation
Main Avenue
Norwalk, Connecticut 06852

John V. Martonchik
Astronomy Department
University of Texas
Austin, Texas 78712

Douglas P. McNutt
U.S. Naval Research Laboratory
Code 7126.2
Washington, D.C. 20390

Lawrence N. Mertz
Smithsonian Astrophysical Observatory
60 Garden St.
Cambridge, Massachusetts 02138

Guy G. Michel
Laboratoire Aimé Cotton
C.N.R.S.
91 Orsay, FRANCE

Richard E. Miller
Michigan State University
Chemistry Department
East Lansing, Michigan 48823

Raymond C. Milward
Société Coderg
15 Impasse Barbier
92 Clichy, FRANCE

Richard I. Mitchell
University of Arizona
Optical Sciences Department
Tucson, Arizona 85721

Michael H. Mosley
Beckman R.I.I.C. Ltd.
Worsley Bridge Rd.
London SE26 ENGLAND

David G. Murcray
University of Denver
Denver, Colorado 80210

Randall E. Murphy
AFCRL (OPR)
Hanscom Field
Bedford, Massachusetts 01730

John S. Nader
National Air Pollution Control Administration
5710 Wooster Pike
Cincinnati Ohio 45227

Douglas J. Neale
Beckman R.I.I.C. Ltd.
Worsley Bridge Rd.
London SE26 ENGLAND

Richard A. Nesbit
Beckman Instruments, Inc.
2500 Harbor Blvd.
Fullerton, California 92634

Ronald Newburgh
AFRCL (LZ)
Hanscom Field
Bedford, Massachusetts 01730

Joseph T. O'Connor
University of California
Space Sciences Laboratory
Berkeley, California 94720

Kenneth H. Olsen
Los Alamos Scientific Laboratory
Los Alamos, New Mexico 87544

David A. Othen
University of Alberta
Chemistry Department
Edmonton 7 Alberta CANADA

W. D. Perkins
Perkin-Elmer Corporation
Norwalk, Connecticut 06852

Clive Perry
Northeastern University
Physics Department
Boston, Massachusetts 02115

Steve Petcherych
University of Saskatchewan
Physics Department
Saskatoon, Saskatchewan

Melvin E. Peterkin
Sun Oil Company
P.O. Box 426
Marcus Hook, Pennsylvania 19061

Jacques Pinard
Laboratoire Aimé Cotton
C.N.R.S.
91 Orsay, FRANCE

Andris I. Priede
Technical Operations
Northwest Industrial Park
Burlington, Massachusetts 01803

Leon Radziemski
Los Alamos Scientific Laboratory
Los Alamos, New Mexico 87544

Charles M. Randall
The Aerospace Corporation
P.O. Box 95085
Los Angeles, California 90045

John W. Reed
AFCRL (OPR)
Hanscom Field
Bedford, Massachusetts 01730

John D. Rex
AFCRL (OPR)
Hanscom Field
Bedford, Massachusetts 01730

Paul L. Richards
University of California
Physics Department
Berkeley, California 94720

John Roche
Perkin-Elmer Corporation
Main Avenue
Norwalk, Connecticut 06852

David H. Rodges
Texas Instruments, Inc.
13500 N. Central Expressway
Dallas, Texas 75222

Ginette F. Roland
Institut d'Astrophysique
B-4200 Cointe-Ougree, BELGIQUE

Harold Romero
University of Missouri
Physics Department
Rolla, Missouri 65401

Arnulf Roseller
Zentralinstitut für Optik und Spektroskopie der DAW
5 Rudower Chaussee
1199 Berlin-Adlershof
DDR

George Rupprecht
Bendix Navigation and Control Division
2796 S. Federal Blvd.
Denver, Colorado 80211

Edgar E. Russell
Santa Barbara Research Center
75 Coromar Drive
Goleta, California 93017

Joel W. Russell
Oakland University
Rochester, Michigan 48019

Hajime Sakai
AFCRL (OPI)
Hanscom Field
Bedford, Massachusetts 01730

Richard B. Sanderson
Ohio State University
Physics Department
Columbus, Ohio 43210

Brian P. Sanford
AFCRL (OPR)
Hanscom Field
Bedford, Massachusetts 01730

Jean-Yves Savard
Laval University
Electrical Engineering Department
Quebec 10, CANADA

William Schade
U.S. Naval Weapons Center
Code 613
Corona, California 91720

Raymond J. Schmitt
McDonnell Douglas Corporation
P.O. Box 516
St. Louis, Missouri 63166

Herbert W. Schnopper
M.I.T.
Center for Space Research
Cambridge, Massachusetts 02139

Ronald L. Schwiesow
ESSA Research Laboratories
Boulder, Colorado 80302

Robert D. Sears
Lockheed Research Laboratories
3251 Hanover St.
Palo Alto, California 94304

Thomas P. Sheahen
Bell Telephone Laboratories
Whippany, New Jersey 07981

Gordon G. Shepherd
York University
Physics Department
4700 Keele St.
Toronto, Ontario, CANADA

George Shipley
University of Colorado
Electrical Engineering Department
Boulder, Colorado 80302

Donald E. Silva
University of California
Applied Physics and Information Sciences
La Jolla, California 92037

George D. Singer
Night Vision Laboratory
Fort Belvoir, Virginia 22060

Edwin Sloane
Time/Data Corporation
490 San Antonio Road
Palo Alto, California 94306

Neil Sloane
Bell Telephone Laboratories
Murray Hill, New Jersey 07974

Glen H. Smerage
Utah State University
Electro-Dynamics Laboratories
Logan, Utah 84321

Donald R. Smith
AFCRL (OPI)
Hanscom Field
Bedford, Massachusetts 01730

Lewis L. Smith
Grumman Aerospace Corporation
Research Department—Plant 35
Bethpage, New York 11714

W. Snellerman
Fysisch Laboratorium
Rijksuniversiteit
Bylhouwerstraat 6
Utrecht, NETHERLANDS

Robert P. Spiers
NASA Langley Research Center
Code 473
Hampton, Virginia 23365

A. T. Stair, Jr.
AFCRL (OPR)
Hanscom Field
Bedford, Massachusetts 01730

Allan J. Steed
Utah State University
Electro-Dynamics Laboratories
Logan, Utah 84321

William H. Steel
C.S.I.R.O.
National Standards Laboratory
City Road
Chippendale
N.S.W. 2008 AUSTRALIA

David L. Steinmetz
University of Arizona
Optical Sciences Center
Tucson, Arizona 85705

John D. Strong
University of Massachusetts
Astronomy Research Facility
Amherst, Massachusetts 01002

Sham Sunder
University of Alberta
Chemistry Department
Edmonton, Alberta, CANADA

Parn Taimsalu
Beckman Instruments International
17 Rue des Pierres-du-Nilon
Geneva, SWITZERLAND

Andrew G. Tescher
General Dynamics
1675 W. Mission
Pomona, California 91766

Leslie W. Thorpe
Beckman R.I.I.C. Ltd.
Worsley Bridge Road
London SE 26 ENGLAND

Paul T. Tomboulian
Oakland University
Rochester, Michigan 48063

Martin Trefler
University of Toronto
Toronto 5 Ontario, CANADA

William F. Ulrich
Beckman Instruments, Inc.
2500 Harbor Blvd.
Fullerton, California 92634

George A. Vanasse
AFCRL (OPI)
Hanscom Field
Bedford, Massachusetts 01730

Reinder J. VanDuinen
University of Groningen
Kapteyn Observatory
Roderi, NETHERLANDS

Thomas J. Walter
AFCRL (OPR)
Hanscom Field
Bedford, Massachusetts 01730

Herbert Walther
Joint Institute for Laboratory Astrophysics
NBS
Boulder, Colorado 80302

David M. Weiss
U.S. Naval Research Laboratory
Washington, D.C. 20390

Ned B. Wheeler
AFCRL (OPR)
Hanscom Field
Bedford, Massachusetts 01730

Pieter R. Wiederhold
Comstock & Wescott, Inc.
765 Concord Avenue
Cambridge, Massachusetts 02138

Geert T. Wijntjes
Block Engineering, Inc.
19 Blackstone St.
Cambridge, Massachusetts 02139

Ronald R. Willey
Willey Optical Corporation
635 W. NASA Blvd.
Melbourne, Florida 22901

Nicholas F. Yannoni
AFCRL (PHF)
Hanscom Field
Bedford, Massachusetts 01730

Hiroshi Yoshinaga
Osaka University
Applied Physics
Osaka, JAPAN

Gene Young
Perkin-Elmer Corporation
Main Avenue
Norwalk, Connecticut 06852

Raymond A. Zickus
Technical Operations, Inc.
South Avenue—Bldg. 3
Burlington, Massachusetts 01803

Harold H. Zwick
York University
4700 Keele St.
Downsview, Ontario, CANADA

END