

THE UNIVERSITY OF MICHIGAN
INDUSTRY PROGRAM
OF THE
COLLEGE OF ENGINEERING

**Interferometric Measurement of
Three-dimensional Temperature Fields**

DONALD W. SWEENEY

September, 1972

IP - 846



ANN ARBOR

ABSTRACT

INTERFEROMETRIC MEASUREMENT OF
THREE-DIMENSIONAL TEMPERATURE FIELDS

by

Donald Wesley Sweeney

Co-Chairmen: Emmett N. Leith
Charles M. Vest

Numerical and analytical techniques are presented which allow asymmetric temperature fields to be reconstructed from the pathlength measurements obtained with holographic interferometry. Analytical reconstruction techniques which have been used in radio astronomy and electron microscopy for a number of years, and recently in interferometry, are presented in the context of interferometric applications in the refractionless limit. These techniques require that optical pathlength data be collected over a 180° angle of view. The required pathlength sampling rate is found to be a function of both the internal structure of the index field and the desired resolution. A direct relationship between the refractive index field and the pathlength function is expressed in terms of the Hilbert transform of the optical pathlength function. An efficient numerical procedure is developed to evaluate the resultant integral equation.

Several numerical schemes are developed which do not require that data be collected over a full 180° angle of view. The angle over which data can be collected is a relative factor; the effective angle of view is proportional to the resolution of the reconstruction and

the physical extent of the refractive index field. The deleterious effect of errors in the pathlength samples increases as the angle of view decreases. All of the reconstruction techniques require redundant data to attain good accuracy. As the angle of view is reduced, the amount of redundant data necessary increases and the resolution attainable along the mean direction of view gets smaller.

Numerical simulations using six different reconstruction techniques show that with a 180° angle of view, all of the schemes provide accurate reconstructions. Examples of simulated reconstructions are presented in which only a 45° angle of view is available, the results reveal increased errors as the angle of view decreases.

The techniques are applied to determine the temperature field in the steady-state convective plume above a short heated horizontal wire in water. Phase grating illumination is used to eliminate the problems associated with coherent speckle. The techniques are also used to reconstruct the temperature field above a heated horizontal plate. The results reveal the same basic characteristics that have been observed in previous flow visualization experiments.

INTERFEROMETRIC MEASUREMENT OF
THREE-DIMENSIONAL TEMPERATURE FIELDS

by
Donald Wesley Sweeney

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mechanical Engineering)
in The University of Michigan
1972

Doctorial Committee:

Professor Emmett N. Leith, Co-Chairman
Associate Professor Charles M. Vest, Co-Chairman
Professor Vedat S. Arpaci
Professor John A. Clark
Dr. Jerry S. Zelenka

ACKNOWLEDGMENTS

I would like to express my gratitude to the members of my dissertation committee. I would particularly like to express my appreciation to Professor Charles Vest who introduced me to the area of coherent optics. Professor Vest's support and guidance contributed substantially to this dissertation. I also would like to thank Professor Emmett Leith for his stimulating lectures and discussions.

I would also like to acknowledge J. L. Walker and J. Zelenka for their contributions in the form of many informative discussions, and Gary Montry for his assistance in the computer graphics. Many thanks are due to Gerry Nowak for typing the dissertation.

This research was performed at the Radar and Optics Division of the Willow Run Laboratories. I am indebted to the staff of the Radar and Optics Division for providing the facilities and advice that made this dissertation possible.

A special note for my wife, Joyce, who spent many lonely evenings alone. Her support and understanding is deeply appreciated.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	ii
LIST OF FIGURES	v
LIST OF TABLES	vii
LIST OF SYMBOLS	viii
CHAPTER 1. Reconstruction of Asymmetric Temperature Fields	1
Introduction	1
Literature Survey	3
CHAPTER 2. Reconstruction by Fourier Transforms	10
Fourier Analysis of Pathlength Data	15
Discrete Input Data	22
Required Sampling Rates	28
CHAPTER 3. Reconstruction by Series Expansion	37
Basic Approach	37
Grid Method	39
Sample Method	45
Frequency Plane Restoration	51
Use of Redundant Data	55
Comparison of Grid and Sample Methods	61
Other Reconstruction Techniques	74
CHAPTER 4. Comparison of the Reconstruction Techniques	82

	Page
CHAPTER 5. Experimental Applications to Temperature Measurement	94
Experimental Considerations	94
Temperature Above a Short Heated Wire	100
Temperature Above a Horizontal Heated Surface	104
CHAPTER 6. Summary and Conclusions	118
Reconstruction by Fourier Analysis and Series Expansions	118
Conclusions	121
REFERENCES	123

LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
2-1	Geometry for the Reconstruction of a Two-Dimensional Refractive Index Field	12
2-2	Notation for Fourier Analysis of the Pathlength Data	16
2-3	Geometric Representation of the Convolution in Eq. (2-16)	21
2-4	Nomenclature for the Discrete Fourier Transform	24
2-5	Representation of the Aliasing Effects Caused by Undersampling	29
2-6	Example of Interpolation Between Known Radial Lines in the Frequency Plane	31
3-1	Nomenclature for the Grid Series Expansion of the Refractive Index Field	41
3-2	Nomenclature for the Sample Series Expansion of the Refractive Index Field	47
3-3	Nomenclature for the Frequency Plane Restoration Technique	52
3-4	Linear System Representation of the Transformation $\overline{Wf} = \overline{\phi}$	62
3-5	Plot of Noise Amplification Factor vs. Angle of View and Resolution of the Reconstruction for 10 Equally Spaced Views and 11 Rays/View	68
4-1	Isometric Plot of Double Gaussian Function	85
4-2	Isometric Plot of Cosine Function	87
4-3	Isometric Plot of Rayleigh Distribution	89
5-1	Plot of Typical Pathlength Function Showing Positions of Center of Fringe and the Possible Effect of Aliasing	97

<u>Figure</u>	<u>Page</u>
5-2 The Effect of Neglecting Refraction, the Path-length of Ray ABC is Assigned to the Hypothetical Straight Ray DEF	99
5-3 Photograph of the Short Heated Wire and Its Support Stand	100
5-4 Schematic of the Experimental System Used in the Heated Wire Experiments, the Object Illumination is Derived from Phase Gratings	101
5-5 Typical View of an Interferogram of the Short Heated Wire in Water	103
5-6 Reconstructed Temperature Field in a Plane Above the Heated Wire	105
5-7 Photograph of the Heated Plate and Its Support Stand	106
5-8 Typical View of the Interferogram of Heated Plate in Water	108
5-9 Isometric Plot of the Reconstructed Temperature Field in a Plane 1 cm Above the Heated Plate (Sample Method, $M = 3$, $N = 3$)	109
5-10 Isothermal Contour Plot for a Plane 1 cm Above the Heated Surface (Sample Method, $M = 3$, $N = 3$)	110
5-11 Isothermal Contour Plot for a Plane 0.5 cm Above the Heated Plate (Sample Method, $M = 3$, $N = 4$)	112
5-12 Sketch of Results of Flow Visualization Obtained by Husar and Sparrow [54] for Free Convection Above Horizontal Heated Surface	114
5-13 Isothermal Contour Plot for Plane 1 cm Above Heated Surface (Grid Method, $M = 3$, $N = 3$)	115
5-14 Isothermal Contour Plot for Plane 0.5 cm Above Heated Surface (Grid Method, $M = 3$, $N = 4$)	116
5-15 Isothermal Contour Plot for Plane 0.5 cm Above Heated Surface (Frequency Plane Restoration, $M = 3$, $N = 4$)	117

LIST OF TABLES

<u>Table</u>		<u>Page</u>
3-1	Comparison of the Sample Method and the Grid Method for Three Truncated Gaussian Functions with 10 Views and 11 Rays/View, $M = 4$ and $N = 4$	71
4-1	Comparison of Results. Input Function: Double Gaussian	86
4-2	Comparison of Results. Input Function: Cosine Function	88
4-3	Comparison of Results. Input Function: Rayleigh Distribution	90
4-4	Comparison of the Sample Method and the Frequency Plane Restoration Technique as a Function of the Mean Directional View for the Double Gaussian Function Shown in Fig. 4-1	93

LIST OF SYMBOLS

a_i	slope of ray as it passes through refractive index field
a_{mn}	coefficients of generating functions used in series expansion of relative refractive index field, see Eq. (3-1)
B_x, B_y	bandwidth in the f_x and f_y directions, respectively
b'_i	y intercept in shifted coordinate system, see Figs. 3-1 and 3-2
c	relaxation parameter used in Eq. (3-60)
E	expected value
e_i	error of temperature measurement at position i defined by Eq. (4-1a)
e_{av}	average error of temperature measurement defined by Eq. (4-1a)
F	Fourier transform of relative refractive index field
f	refractive index relative to ambient ($n - n_o$)
$\bar{f}(x, y)$	series expansion of relative refractive index field
f_j, \bar{f}	column vector of the coefficient of the series expansion of the relative refractive index
f_x, f_y	coordinates in the Fourier transform plane
$f_j^{(k)}$	coefficients of the series expansion for the relative refractive index after the k th iteration, defined in Eqs. (3-57) or (3-60)
G_{mn}	Fourier series coefficients defined by Eq. (3-24)
G_j	column vector of the elements of G_{mn}
H_{mn}	generating functions used in Eq. (3-1)
i	$\sqrt{-1}$

J_n	n th order Bessel function of the first kind
L_x, L_y	linear extent of the relative refractive index in the x and y direction, respectively (see Fig. 2-4)
l_x, l_y	width of the grid elements in the x and y directions, as shown in Fig. 3-1, or the spacing of the sinc function in the x and y direction, as shown in Fig. 2-2
M, N	number of terms used in the series expansion of the relative refractive index field in the x and y direction, respectively
m,n,l,k	integers used in the discrete Fourier transform (see Eqs. (2-17) to (2-20))
n	refractive index
n_0	ambient refractive index
p	peak-to-peak phase delay of phase grating
q'_i	dummy variable equal to $-b'_i/a_i$ in Eq. (3-17)
R	polar position vector in the Fourier transform plane, or noise amplification factor defined by Eq. (3-36)
$R^{(k)}$	residual vector after k th iteration in ART procedure, defined by Eq. (3-58)
r	polar coordinate in the spatial plane equal to $(x^2 + y^2)^{1/2}$
\bar{r}	position vector used in Eq. (2-2)
rect(x)	rectangular function defined as $\begin{cases} 1 & \text{if } x \leq 1/2 \\ 0 & \text{if } x > 1/2 \end{cases}$
S_i, S	path of light ray as it traverses the refractive index field
SBWP	space-bandwidth product defined by Eq. (2-23)
(S/N)	signal-to-noise ratio defined by Eq. (3-34)

$\text{Sgn}(x)$	sign function defined as $\begin{cases} +1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$
$\text{sinc}(x)$	sinc function defined as $\sin(\pi x)/\pi x$
t	matrix transpose
\bar{U}, \bar{V}	orthonormalized matrices used to transform W_{ij} , described after Eq. (3-37)
U	optical amplitude
v, v'	rotated coordinate system shown in Fig. 2-2
W	see Fig. 2-5
$W(a_i, \rho_i)$	coefficients of the matrix generated by integrating the series expansion of $f(x, y)$ along a ray with slope a_i and radial position ρ_i
W_{ij}, \bar{W}	matrix notation for $W(a_i, \rho_i)$
x, y	spatial coordinates
x', y'	shifted spatial coordinates shown in Figs. 2-1 and 2-2
z	dummy variable equal to $r \sin(\phi - \theta)$ in Eq. (2-14) and equal to $-b'_i/a_i$ in Eq. (3-8)
$\mathcal{F}, \mathcal{F}^{-1}$	Fourier transform operators defined by Eqs. (2-4) and (2-5)
α_i	summation of the squares of the elements in the i th row of the matrix W_{ij} , defined by Eq. (3-57)
β	angle defined in Fig. 2-2
$\bar{\nabla}$	gradient operator
$\delta(x)$	Dirac δ function
ϵ	arbitrary small number used, for example, in Eq. (3-60)
θ	angular orientation of ray as shown in Fig. 2-1

θ_{eff}	effective angle of view defined by Eq. (3-11)
Λ	diagonal matrix consisting of the eigenvalues of the inner product (W_{ji}, W_{ij})
λ	wavelength of light
λ_i	singular values of the matrix W_{ij} or Lagrange multipliers used in Eq. (3-50)
ρ	radial position of ray shown in Fig. 2-1
$\bar{\Sigma}$	diagonal matrix consisting of the singular values of the matrix W_{ij}
σ^2	variance of the additive noise defined in Eq. (3-43)
ϕ_i, ϕ	pathlength function defined by Eq. (2-1)
$\phi(\rho, \theta)$	pathlength function defined by Eq. (2-3)
$\bar{\phi}$	derivative of pathlength function as shown in Fig. 2-3
ϕ	polar coordinate in spatial plane

CHAPTER 1

RECONSTRUCTION OF ASYMMETRIC TEMPERATURE FIELDS

Introduction

There are a number of optical methods used in experimental heat transfer to measure temperature. Among them are schlieren, shadow-graph, and interferometric methods. Of particular interest here is optical interferometry which utilizes the dependence of the refractive index on the temperature of an optically transparent medium to reveal the temperature field. This dissertation deals with the applications of interferometry to the measurement of three-dimensional temperature fields in transparent media. Several analytical and numerical procedures are developed which allow three-dimensional temperature fields to be reconstructed from experimental data obtained using holographic interferometry. The amount of data necessary and the resolution in the reconstructed temperature field will be discussed in detail. The reconstruction techniques will be compared in a numerical simulation and applied to actual temperature measurement experiments.

The information available from an interferogram is in the form of alternate bright and dark bands called fringes. The fringes indicate contour lines of constant optical pathlength. The pathlength is equal to the integral of the refractive index field along the ray paths. Since it is the value of this integral that is available from the interferogram, the three-dimensional temperature field can be

obtained by inverting an integral equation. Specifically, the integral equation to be solved is

$$\int_{S_i} f(x, y, z) dS_i = \phi_i \quad (1-1)$$

where $f(x, y, z) = n(x, y, z) - n_0$.

Here the desired quantity, $f(x, y, z)$, is the refractive index relative to some reference n_0 , ϕ_i is the pathlength function available from the interferogram, and the line integral is evaluated along the path S_i of a light ray which traverses the refractive index field $n(x, y, z)$. When Eq. (1-1) is solved for the three-dimensional refractive index distribution $[f(x, y, z) + n_0]$, the temperature field can be obtained using the relationship between refractive index and temperature.

As will be shown, the general solution of Eq. (1-1) can be obtained only if the pathlength function, ϕ_i , is available for ray paths through the refractive index field oriented in different angular positions. Until recently, the general solution to Eq. (1-1) had received little attention because classical interferometers, which use plane-wave illumination, only provided information about the integral of the refractive index field for a single direction of illumination. The solution to Eq. (1-1) could only be obtained if the pathlength data obtained from the single projection available provided essentially all of the information necessary to unambiguously

determine the refractive index distribution. This is the case if the index field is two-dimensional or axially symmetric.

Holographic interferometry does not share this limitation. The high information content of holograms makes it possible to record interferograms with multidirectional or even diffuse illumination of the transparent object being studied. A single holographic interferogram can record a measure of the optical pathlength for rays passing through the object in many different directions. The methods discussed and developed in this dissertation, coupled with the use of holographic interferometry, provide a technique for reconstructing three-dimensional temperature fields from multidirectional interferometric pathlength data.

The details of optical interferometry will not be discussed here. For a detailed summary of the principles involved and applications to heat transfer, the reader is referred to references 1 through 5; however, a brief description of the procedures involved in holographic interferometry is given in Chapter 5.

Literature Survey

The integral equation of three-dimensional interferometry, Eq. (1-1), appears in different physical contexts in several other areas of science. In particular, these areas are radio astronomy, plasma physics, x-ray photography, and electron microscopy. In these other areas, the line integral of the three-dimensional scalar field, represented here as $f(x, y, z)$, is evaluated along straight lines. In Chapter 2, an approximation will be introduced so that the path

of integration in Eq. (1-1) is also evaluated along straight lines here; therefore, the literature in these other areas is relevant to the present discussion and will be included here.

Bracewell [6], working in the field of radio astronomy, was apparently the first to present a detailed discussion of the solution of Eq. (1-1) using Fourier analysis. He showed that the radio signal received by a long narrow synthetic aperture radio antenna could be used to determine the values along a properly oriented line of the two-dimensional Fourier transform of the radio source. By rotating the synthetic aperture, the complete two-dimensional radio source could be reconstructed. Bracewell [6] also showed the close relationship between the solution of Eq. (1-1) and the Abel integral. Implementation of this early work by Bracewell [6] was hampered by the lack of sophisticated computing machines. In fact, the procedures presented were designed for a desk calculator. Later, Bracewell and Wild [7] developed a scheme which simplified many of the numerical complexities involved. An example of the application of these techniques is presented by Taylor [8] who utilized the analysis of Bracewell [6] to reconstruct the radio galaxy 3C192.

Several years earlier, O'Brien [9] had suggested the possibility of reconstructing a two-dimensional radio source from the signals received by a long slender antenna but did not present any formal techniques for reconstructing a general asymmetric distribution. Smead and Wild [10] generalized O'Brien's [9] work and presented a technique to reconstruct asymmetric distributions which were basically the same as Bracewell's [6] procedure. They discussed the effects of

using only a finite number of angular orientations of the antenna, and showed the loss of resolution that results from using an antenna of finite width.

De Rosier and Klug [11] noticed that Eq. (1-1) appears in electron microscopy and independently developed the same result as Bracewell [6]. The work of De Rosier and co-workers at Cambridge has received considerable attention because it provides a technique for reconstructing the internal structure of individual biological structures from electron micrographs. Klug [12] presents an excellent summary of the work done in this area and some of the problems encountered in practical applications. Crowther, et al. [13] discuss several numerical procedures which can be used to implement the reconstruction schemes. They determined that to obtain the high resolution necessary in electron microscopy, data must be collected over a full 180° . An example of the many experimental applications of these techniques is presented by Crowther [14], who reconstructs a negatively stained tomato bushy stunt virus.

Hoppe [15] noted the relationship between optical superresolution of finite objects as presented by Harris [16] and the three-dimensional reconstruction procedures. The superresolution procedures presented the possibility of reconstructing an object when a finite number of projections were available which cover only a restricted region in the Fourier transform space. He apparently did not apply the technique experimentally.

Tretiak [17] and Tretiak, et al. [10] applied the reconstruction techniques developed by Bracewell [6] to determine the internal structure of a femur from x-ray shadowgraphs. Their results clearly show the internal density variations within the bone.

Rowley [19], working in the area of interferometry, developed the Fourier transform relationship between the pathlength measurements obtained from interferograms and the refractive index field. Although the results of the analysis are the same as those obtained by Bracewell [6], this paper is noteworthy because it represents the first analytical work done in the area of three-dimensional interferometry.

Berry and Gibbs [20] presented an analysis which directly gives $f(x, y)$ as a function of ϕ in Eq. (1-1). Their work was concerned with evaluation of x-ray shadowgraphs. The reconstruction of $f(x, y)$ does not directly require the use of Fourier transforms. The procedure does require that the principal value of an integral be determined. They did not present a numerical procedure to evaluate the integral. Tanner [21] discussed the direct applicability of the results of Berry and Gibbs to interferometry. Although he did not numerically evaluate the integral, he did present arguments to show that the principal value of the integral is finite. Alwang, et al. [22] reported that A. Pal independently derived the Berry and Gibbs [20] solution. To avoid the singularity, he introduced an approximate representation of the integrand; it contains a "resolution parameter" the optimum value of which is dependent upon the structure of the field being studied. This method was used to measure the density

distribution in a slot flame. Junninger and van Haeringer [23] used the analysis of Rowley [19] to derive the same direct relationship between $f(x, y)$ and ϕ in Eq. (1-1); they did present a numerical procedure to evaluate the principal value of the integral using a Taylor series expansion.

Ramachandran [24] and Ramachandran and Lakshminarayanan [25] have developed a numerical convolution procedure to reconstruct $f(x, y)$ from ϕ directly. These procedures can be interpreted as a discrete form of the Berry and Gibbs [20] integral solution.

Several numerical procedures have been developed to solve Eq. (1-1) using a series expansion for $f(x, y)$. The available data ϕ is used to determine the coefficients of the series expansion. Maldonado [26] derived an orthogonal series which is invariant to a rotation of axis; this series was developed using the procedures suggested by Bhatia and Wolf [27]. Maldonado and Olsen [28] used the series to develop a scheme to reconstruct the local emission coefficients of a radiating plasma. The coefficients were determined using the orthogonality condition.

Matulka and Collins [29] applied the series of Maldonado [26] to interferometry. They used this procedure to reconstruct the density field in an air jet which was slightly tilted with respect to the optical axis in order to simulate an asymmetric density distribution. The details of the computer algorithms used are given by Matulka in Ref. 30.

Alwang, et al. [22] discussed a discrete element technique in which the phase object is considered to be divided into N finite grid elements, each of which has a uniform refractive index. This technique was found to be extremely sensitive to factors such as grid size. Although little detail regarding this approach was presented, it was apparently not too satisfactory.

A group of researchers at State University of New York at Buffalo have developed several iterative techniques which they call Algebraic Reconstruction Techniques (ART) [31-33]. Their techniques are intended to be an alternative approach to the methods developed by De Rosier and Klug [11] to determine the internal structure of biological specimens from electron micrographs. Gordon, et al. [31] presented two iterative procedures to reconstruct $f(x, y)$. The procedure called the additive technique seems to produce the best results. Bender, et al. [32] applied these iterative techniques to reconstruct the internal structure of ribosomes. Herman and Rowland [33] have shown several examples of the resolution attainable using the iterative techniques.

The preceding references review the existing research related to the solution of Eq. (1-1). In summary, all of the techniques experimentally applied to date have used the data collected over a 180° field of view. Two authors, Hoppe [15] and Crowther [13], have suggested that, in theory, it is possible to use a smaller angle of view, but neither have applied procedures which realize this result. While the various solution techniques come from a variety of different areas of application, there have been only a few relevant publications

in the area of interferometry. Only two papers in interferometry, Matulka and Collins [30] and Alwang [23], present results of actual experimental reconstructions and both of these require a 180° field of view.

References which deal with the holographic techniques that will be applied to experimentally collect the pathlength data are not included in this section but will be referenced in the text where appropriate.

CHAPTER 2

RECONSTRUCTION BY FOURIER TRANSFORMS

Repeating from Chapter 1, the fundamental integral equation to be solved in three-dimensional interferometry is

$$\int_{S_i} f(x, y, z) dS_i = \phi_i \quad (2-1)$$

where $f(x, y, z) = n(x, y, z) - n_0$; $f(x, y, z)$ is the refractive index relative to a known value n_0 , S_i is the distance along the i th ray traversing the test section, ϕ_i is the known pathlength function, and the integral is evaluated between boundaries imposed by the experimental configuration or by a priori knowledge regarding $f(x, y, z)$. The limits of integration in Eq. (2-1) can be arbitrarily extended, however, since the integrand is zero outside the boundaries of $f(x, y, z)$. In mathematical terms, the objective is to reconstruct a real scalar function--the relative refractive index field--from the scalar values of line integrals through the field--the optical pathlengths--where the values of the line integral are available for many different positions and orientations, and the limits of integration are established by the boundaries of the scalar function.

As is well known (cf. [5]), the path of a ray as it traverses an inhomogeneous refractive index field is a curved path which satisfies the eikonal equation,

$$\frac{d}{dS} \left(n \frac{d\bar{r}}{dS} \right) = \bar{\nabla} n \quad (2-2)$$

where \bar{r} is a position vector. Thus, the path of integration depends on the unknown refractive index field. The general solution to the coupled nonlinear equation set (2-1) and (2-2) is unknown. It is not clear that the pathlength measurements obtained from interferometry will provide all the information necessary to obtain a solution since the position of ray incidence is unknown.

The departure of a light ray from a straight line is proportional to the gradient of the refractive index field. For many problems of interest encountered in interferometry, the gradient of the refractive index variations is small enough to produce negligible ray curvature. In this refractionless limit, Eq. (2-1) alone is the basic equation, where the line integrals are evaluated along straight lines.

Unless noted otherwise, ray curvature due to refraction will be neglected for the remainder of the dissertation.

If refraction is neglected, it is possible to simplify the presentation of three-dimensional interferometry by suppressing the third dimension. Any solution for the two-dimensional (2-D) problem depicted in Fig. 2-1 can be easily generated to include the third dimension. Or, equivalently, the 2-D problem can be solved for planes distributed along the third dimension and the complete 3-D distribution "built-up" in layers. In this modified form, the relative refractive index field is a function of only two spatial coordinates, x and y , where only those rays in the x - y plane are included. This simplification does not change the basic nature of the problem and is only done to simplify the presentation.

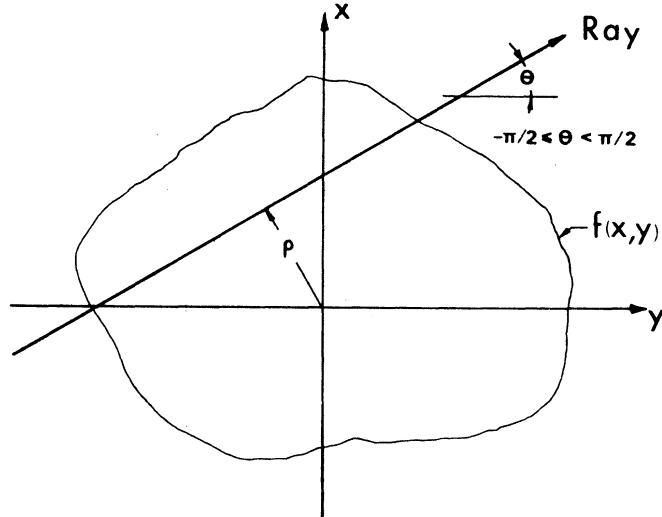


Fig. 2-1. Geometry for the Reconstruction of a Two-Dimensional Refractive Index Field. A Ray is Specified by the Parameters ρ and θ .

Various rays may be specified by their polar coordinates ρ and θ (Fig. 2-1). Equation (2-1) becomes

$$\int_S f(x, y) dS = \phi(\rho, \theta) \quad (2-3)$$

where ρ can vary to cover the domain of $f(x, y)$ and θ can vary between $\pm\pi/2$. Due to experimental conditions, however, the pathlength data

may not be available over a full 180° range. The total variation in θ may be limited by obstructions by parts of the experimental system, or the experimental system necessary to realize a complete 180° range of view may be prohibitively complex. The total angular extent over which values of θ are available will be referred to as the total angle of view or, more simply, as the angle of view. Unless otherwise noted, the angle of view will be centered about the x axis.

If the refractive index field is radially symmetric or if the refractive index field is independent of one of the coordinates, then the solution of Eq. (2-3) is known and has been used extensively in interferometric studies.

When the refractive index field is independent of one of the coordinates, only one view is required if the view is normal to the remaining coordinate axis. For example, if $f(x, y) = f(y)$ and the pathlengths are collected for $\theta = 0$, and the extent of the relative refractive index field is L in the x direction, then Eq. (2-3) becomes

$$\phi(y) = \int_{-L/2}^{+L/2} n(y) dx = Ln(y)$$

The solution is

$$n(y) = \phi(y)/L$$

For the radially symmetric case, the refractive index depends only on r where $r = x^2 + y^2$. So that the pathlength measurements $\phi(\rho, \theta)$ will be the same for all θ and only one direction of view is necessary. If the single required view is taken with $\theta = 0$ in Fig. 2-1 then, after extending the limits of integration, Eq. (2-3) can be written as

$$\phi(x) = 2 \int_{r=x}^{\infty} f(r) dy = 2 \int_x^{\infty} \frac{f(r)r dr}{(r^2 - x^2)^{1/2}}$$

This is a form of Abel's integral equation. The solution may be obtained by converting the integral into a convolution integral by using the appropriate coordinate transformations. The solution for $f(r)$ is

$$f(r) = \frac{-1}{\pi} \int_r^{\infty} (x^2 - r^2)^{1/2} \frac{d}{dx} \left[\frac{\frac{\partial \phi(x)}{\partial x}}{x} \right] dx$$

There are several papers in the literature [6,10,11,20], primarily from areas of application other than interferometry, which present the analytical relationships necessary to reconstruct an asymmetric refractive index field. The basic solution will be presented here and a fast numerical procedure to apply the analytical solution to real problems with discrete amounts of data will be developed.

Throughout this dissertation, Fourier transforms will be used extensively in the analysis and interpretation of the results. The

Fourier transforms will typically be a function of two spatial variables. A brief introduction to the nomenclature of the two-dimensional Fourier transform will be presented here; for a complete description of the concepts involved, refer to Refs. 34 and 35.

The Fourier transform of a complex function f of two independent variables x and y , represented as $\mathcal{F}(f)$, is defined by

$$\mathcal{F}(f) = F(f_x, f_y) = \iint_{-\infty}^{+\infty} f(x, y) \exp\left[-i2\pi(f_x x + f_y y)\right] dx dy \quad (2-4)$$

The transform is a complex function of two independent variables f_x and f_y , commonly referred to as frequencies. The inverse Fourier transform, represented as $\mathcal{F}^{-1}(F)$, is defined by

$$\mathcal{F}^{-1}(F) = f(x, y) = \iint_{-\infty}^{+\infty} F(f_x, f_y) \exp\left[i2\pi(f_x x + f_y y)\right] df_x df_y \quad (2-5)$$

The Fourier transform may be regarded as a decomposition of the function f into a linear combination of periodic kernel functions of the form $\exp\left[i2\pi(f_x x + f_y y)\right]$ with weighting functions $F(f_x, f_y)$. For a particular frequency pair (f_x, f_y) , the kernel function is directed at an angle of $\tan^{-1}(f_y/f_x)$ with respect to the x axis and has a frequency of $\left(f_x^2 + f_y^2\right)^{1/2}$.

Fourier Analysis of Pathlength Data

Using the concept of the Fourier transform as a linear decomposition, it is possible to derive the relationships necessary to invert Eq. (2-3). Repeating Eq. (2-3) for reference, the basic

equation to be solved is

$$\int_S f(x, y) dS = \phi(\rho, \theta) \quad (2-3)$$

where the notation is shown in Fig. 2-2. Let $F(f_x, f_y)$ represent the two-dimensional Fourier transform of $f(x, y)$. Clearly, if F can be reconstructed from the pathlength data, then f can be obtained also by inverse transforming F (Eq. 2-5).

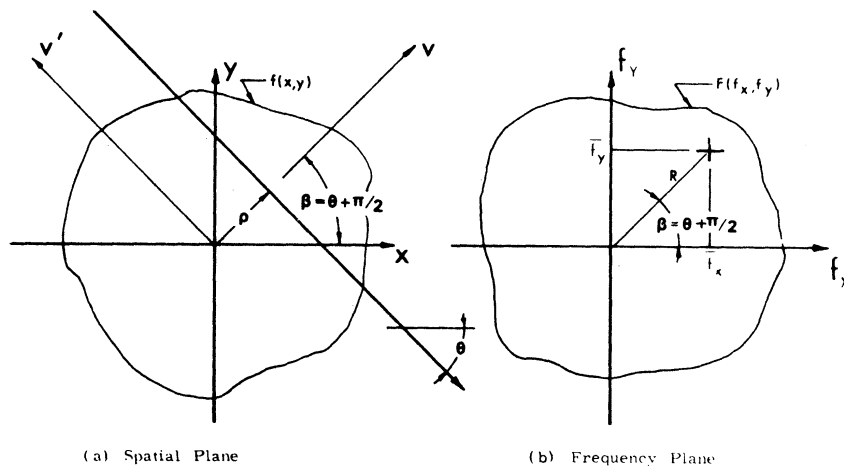


Fig. 2-2. Notation for Fourier Analysis of the Pathlength Data

Consider a point in the frequency plane of f located at (\bar{f}_x, \bar{f}_y) ; the total frequency is $R = (\bar{f}_x^2 + \bar{f}_y^2)^{1/2}$, the kernel function represented by this frequency pair is directed at an angle β with respect to the x axis, where $\beta = \tan^{-1} (\bar{f}_y/\bar{f}_x)$. By definition, the value of F at the point R is

$$F(\bar{f}_x, \bar{f}_y) = \iint_{-\infty}^{+\infty} f(x, y) \exp \left[-i2\pi(\bar{f}_x x + \bar{f}_y y) \right] dx dy \quad (2-6)$$

For convenience, consider a coordinate system (v, v') with the v axis at an angle β with respect to the x axis, as shown in Fig. 2-2a. Equation (2-6) becomes, after some rearranging,

$$F(\bar{f}_x, \bar{f}_y) = \int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{+\infty} f \left[x(v, v'), y(v, v') \right] dv' \right\} \exp -i2\pi(Rv) dv \quad (2-7)$$

The bracketed integral in Eq. (2-7) is a function of v alone; the value of the integral is equal to the integral of $f(v, v')$ along a line parallel to the v' axis. From Fig. 2-2a and the definition of the pathlength, it is clear that

$$\int_{-\infty}^{+\infty} f(v, v') dv' \equiv \phi(\rho, \beta - \pi/2) \equiv \phi(\rho, \theta) \quad (2-8)$$

Equation (2-7) becomes, noting that $v \equiv \rho$,

$$F(\bar{f}_x, \bar{f}_y) = \int_{-\infty}^{+\infty} \phi(\rho, \theta) \exp \left[-i2\pi(R\rho) \right] d\rho \quad (2-9)$$

Since R and β were arbitrarily selected, they may be considered to be variables rather than fixed parameters. If the two polar coordinates, R and θ , are used as the coordinates in the frequency plane, Eq. (2-9) may be written

$$F(R, \beta) = F\left(R, \theta + \frac{\pi}{2}\right) = \int_{-\infty}^{+\infty} \phi(\rho, \theta) \exp\left[-i2\pi(R\rho)\right] d\rho \quad (2-10)$$

Equation (2-10) represents the basic relationship between the known pathlengths and the Fourier transform of the refractive index field. It shows that the Fourier transform of the pathlength data collected for one direction of view determines one radial line through the origin in the frequency plane of $f(x, y)$.

Recently, Berry and Gibbs [20], Alwang, et al. [22], and Junginger and van Haeringen [23] have independently shown that Eq. (2-10) can be formally inverted to obtain a direct relationship between the pathlength data and $f(x, y)$.

The Fourier transform defined in Eq. (2-4) can be expressed in polar coordinates by using the coordinate transformations

$$\begin{aligned} x &= r \cos \phi \\ y &= r \sin \phi \\ f_x &= R \cos \beta = -R \sin \theta \\ f_y &= R \sin \beta = R \cos \theta \end{aligned} \quad (2-11)$$

Equation (2-5) becomes

$$f(r, \phi) = \int_{-\pi/2}^{+\pi/2} d\theta \int_{-\infty}^{+\infty} dR |R| F(R, \theta) \exp \left[+i2\pi Rr \cdot \sin(\phi - \theta) \right] \quad (2-12)$$

where the symmetry properties of the transform of a real function have been used. $|R|$ can be written as

$$|R| \equiv \frac{(-i2\pi R) \cdot [i\pi \text{Sgn}(R)]}{2\pi^2} \quad (2-13)$$

where

$$\text{Sgn}(R) = \begin{cases} +1 & \text{if } R > 0 \\ 0 & \text{if } R = 0 \\ -1 & \text{if } R < 0 \end{cases}$$

If $z = r \sin(\phi - \theta)$, then Eq. (2-12) becomes

$$f(r, \phi) = \frac{1}{2\pi^2} \int_{-\pi/2}^{+\pi/2} d\theta \int_{-\infty}^{+\infty} dR (-i2\pi R) [i\pi \text{Sgn}(R)] \cdot F(R, \theta) \exp(i2\pi Rz) \quad (2-14)$$

Using the definition of the Fourier transform, a few fundamental transform pairs, and Eq. (2-10), it is possible to rewrite Eq.

(2-14) as

$$f(r, \phi) = \frac{1}{2\pi^2} \int_{-\pi/2}^{+\pi/2} d\theta \left[\frac{\partial \phi(z, \theta)}{\partial z} * \frac{1}{z} \right] \quad (2-15)$$

where the "*" denotes the convolution. Expanding the convolution integral and replacing z with $r \sin(\phi - \theta)$, the final result for $f(r, \phi)$ is

$$f(r, \phi) = \frac{1}{2\pi^2} \int_{-\pi/2}^{+\pi/2} d\theta \int_{-\infty}^{+\infty} \frac{\partial\phi(\rho, \theta)/\partial\rho}{r \sin(\phi - \theta) - \rho} d\rho \quad (2-16)$$

This is the basic relationship between the pathlength function and the refractive index field.

Equation (2-15) provides an informative interpretation of the result. The integral over θ is a linear summation of the contributions for each direction of view. The contribution from each value of θ is equal to the convolution of the derivative of the pathlength function with $(1/z)$ evaluated at $z = r \sin(\phi - \theta)$. Figure 2-3 shows a graphical representation of the convolution. The value of the contribution is equal to the net area of the cross-hatched region in Fig. 2-3.

The integrand of Eq. (2-16) becomes singular at $\rho = r \sin(\phi - \theta)$; this will present no numerical difficulties for any pathlength function encountered in real applications. Since $(1/z)$ is an odd function, the contributions to the integral near the origin tend to cancel and the integral remains finite. Let

$$\bar{\phi}(z, \theta) = \partial\phi(z, \theta)/\partial z$$

as shown in Fig. 2-3. The net contributions to the integrand of Eq. (2-16) from points located symmetrically about the origin will

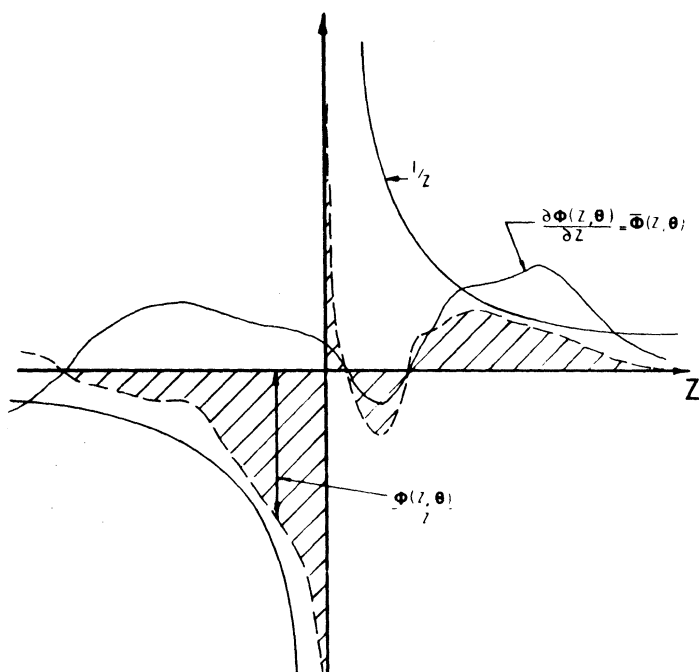


Fig. 2-3. Geometric Representation of the Convolution in Eq. (2-16)

be no larger than the value of the convolution

$$\begin{aligned}
 \bar{\phi}(z, \theta) * [\delta(z - \epsilon) - \delta(z + \epsilon)] & \\
 &= \mathcal{F}^{-1} \left\{ \mathcal{F}[\bar{\phi}(z, \theta)] \left(e^{-i2\pi f_z \epsilon} - e^{i2\pi f_z \epsilon} \right) \right\} \\
 &= \mathcal{F}^{-1} \left\{ 2\epsilon(-2i\pi f_z) \mathcal{F}[\bar{\phi}(z, \theta)] \right\} \quad \text{for } (\epsilon \ll 1) \\
 &= 2\epsilon \frac{\partial \bar{\phi}(z, \theta)}{\partial z}
 \end{aligned}$$

As ϵ tends to zero, the net contribution to the integrand also tends to zero if $\partial^2 \phi(z, \theta) / \partial z^2$ is finite. A requirement which will always be met in real applications. The integral over ρ in Eq. (2-16) is the Hilbert transform of $\phi(\rho, \theta)$. In the next section on discrete data, it will be shown that the properties of Hilbert transforms, along with the use of discrete data, guarantee the boundedness of Eq. (2-16).

The contribution to the value of the integrand of Eq. (2-16) from rays which pass through the refractive index field are damped-out at a rate proportional to $(1/z)$, where z is the radial distance from the point (r, ϕ) . Thus, it may be possible to choose some limit for z beyond which the contributions to the convolution from the pathlength data will be negligible. This procedure of using only those rays which pass within a fixed distance from the point to be reconstructed was simulated numerically. Accurate answers could be obtained, but the accuracy was found to be a strong function of the radial distance used, the particular point under consideration, and the input distribution.

Discrete Input Data

In theory, pathlength data can be collected for all values of θ and the transform plane built-up, a line at a time, until the frequency plane is completely filled. In practice, however, data can only be collected for a finite number of views and positions.

When discrete data is used, the Fourier transforms defined in Eqs. (2-4) and (2-5) must be modified to accept discrete data; the resultant equations are called discrete Fourier transforms. A brief introduction to the nomenclature of the discrete Fourier transforms is presented here; for a complete description of the concepts involved, refer to Refs. 36 and 37.

The discrete Fourier transform of a two-dimensional function $f(x, y)$ is defined as

$$F\left(\frac{m}{L_x}, \frac{n}{L_y}\right) = \frac{1}{2B_x 2B_y} \sum_{\ell=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} f\left(\frac{\ell}{2B_x} \cdot \frac{k}{2B_y}\right) \exp\left[-i2\pi\left(\frac{m\ell}{2B_x L_x} + \frac{nk}{2B_y L_y}\right)\right] \quad (2-17)$$

The discrete inverse Fourier transform is defined as

$$f\left(\frac{\ell}{2B_x} \cdot \frac{k}{2B_y}\right) = \frac{1}{L_x L_y} \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} F\left(\frac{m}{L_x} \cdot \frac{n}{L_y}\right) \exp\left[i2\pi\left(\frac{m\ell}{2B_x L_x} + \frac{nk}{2B_y L_y}\right)\right] \quad (2-18)$$

where L_x and L_y are the spatial extents of the function $f(x, y)$ in the spatial plane, and B_x and B_y are the bandwidths of the Fourier transforms, as shown in Fig. 2-4. Values of the function f located between the sampled locations $x = \ell/2B_x$ and $y = k/2B_y$ are obtained by expanding the series

$$f(x, y) = \sum_{l=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} f\left(\frac{l}{2B_x}, \frac{k}{2B_y}\right) \cdot \text{sinc} \left[\frac{l}{2B_x} \left(x - \frac{l}{2B_x} \right) \right] \cdot \text{sinc} \left[\frac{k}{2B_y} \left(y - \frac{k}{2B_y} \right) \right] \quad (2-19)$$

where

$$\text{sinc } x \equiv \frac{\sin \pi x}{x}$$

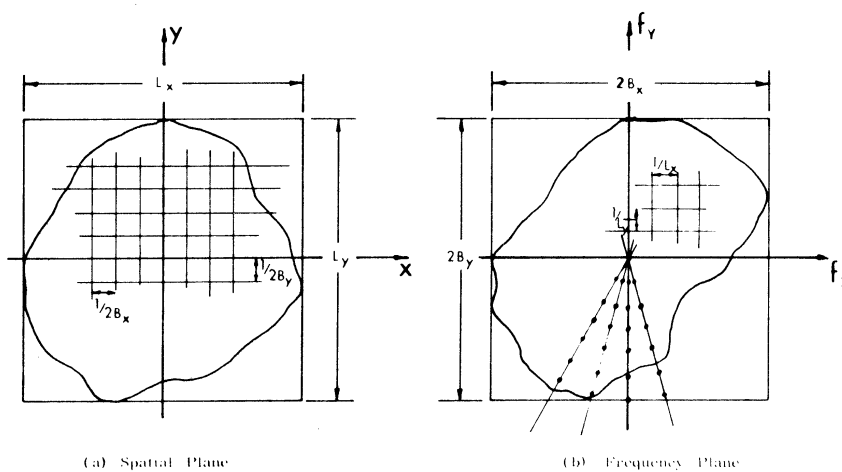


Fig. 2-4. Nomenclature for the Discrete Fourier Transform

A similar expansion can be used in the frequency plane,

$$F(f_x, f_y) = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} F\left(\frac{m}{L_x}, \frac{n}{L_y}\right) \operatorname{sinc} \left[\frac{1}{L_x} \left(x - \frac{m}{L_x} \right) \right] \cdot \operatorname{sinc} \left[\frac{1}{L_y} \left(y - \frac{n}{L_y} \right) \right] \quad (2-20)$$

One procedure which can be used to reconstruct the Fourier transform of the refractive index field is to use a discrete form of Eq. (2-10). As pathlength data is taken for different angular orientations, the transform plane is determined at sampled locations along radial lines extending from the origin. Figure 2-4b shows several typical positions at which the transform plane is determined. The two-dimensional inverse Fourier transform, Eq. (2-18), and the interpolation series, Eq. (2-19), can be evaluated to determine $f(x, y)$. Before the inverse transformation can be used, however, values in the frequency plane must be specified at the sample points required by Eq. (2-18); i.e., values must be specified at the intersection points of a rectangular grid overlapping the frequency plane; a portion of this grid is shown in Fig. 2-4b.

There are two ways in which the frequency plane may be "reformatted." One technique is to apply the concepts developed by J. L. Harris [16] in superresolution; this technique, although mathematically rigorous, requires solving a prohibitatively large set of simultaneous algebraic equations. (Refer to Chapter 3 for a discussion of this method.) Another method is to use linear interpolation

to determine the values at the points on the grid from the known values on the adjacent radial lines. The success of this technique depends on the accuracy of the linear interpolation. The work of J. L. Harris [16] is again relevant. The result, which will be discussed in the next section, is that linear interpolation may be used if the pathlength data is taken for a sufficiently large number of angular orientations.

A direct procedure to reconstruct the refractive index field from the pathlength data and avoid the problems associated with reformatting the frequency plane, is to apply Eq. (2-16). Although it was shown that the principal value of the integral in Eq. (2-16) is finite, numerical evaluation of the integral requires some care.

Equation (2-16) is repeated here for reference

$$f(r, \phi) = \frac{1}{2\pi^2} \int_{-\pi/2}^{+\pi/2} d\theta \int_{-\infty}^{+\infty} \frac{\partial\phi(\rho, \theta)/\partial\rho}{r \sin(\phi - \theta) - \rho} d\rho \quad (2-16)$$

Let $z = r \sin(\phi - \theta)$; then using the convolution theorem and some basic properties of the Fourier transform, it is easy to show

$$\begin{aligned} \int_{-\infty}^{+\infty} \frac{\partial\phi(\rho, \theta)/\partial\rho}{r \sin(\phi - \theta) - \rho} d\rho &= \left. \frac{\partial\phi(z, \theta)}{\partial z} * \frac{1}{z} \right|_{z = r \sin(\phi - \theta)} \\ &= 2\pi^2 \mathcal{F}^{-1} \left\{ |f_z| \mathcal{F}[\phi(z, \theta)] \right\} \Big|_{z = r \sin(\theta - \phi)} \end{aligned} \quad (2-21)$$

For a fixed θ , the value of the integral over ρ in Eq. (2-16) may be evaluated by taking the Fourier transform of the pathlength data, multiplying by a linear phase shift, and inverse transforming. This operation determines the value of the integral for all values of r . The procedure is repeated for each view.

This technique is particularly attractive when discrete data is used. The fast Fourier transform algorithms developed by Cooley and Tukey [38] require only $2N \ln N$ computer operations to compute an N -point discrete Fourier transform. If W views are used, then a table of values of the inner integral in Eq. (2-16) for various r and θ can be obtained in only $W(4N \ln N + N)$ computer operations--a rather remarkable result. For typical computers available today and reasonable values for N and W , the total computer time required to complete the table is less than one second. The value of the refractive index at any point is obtained by combining the proper W entries in the table. Values of the integral which lie between two tabulated values can be obtained by interpolation.

The integral over ρ in Eq. (2-16) is the Hilbert transform of $\partial\phi(\rho, \theta)/\partial\rho$ evaluated at $r \sin(\phi - \theta)$. It is easy to show, using a few properties of the Hilbert transform, that if $\phi(\rho, \theta)$ is sampled at a rate equal to twice the spatial bandwidth and all of the samples have finite values, then the integral over ρ in Eq. (2-16) is finite, i.e., convergent. These two requirements on the sampled values of $\phi(\rho, \theta)$ will always be met in real applications. The Hilbert transform $\hat{g}(u)$, of a function $g(x)$ is defined as

$$\hat{g}(u) = \int_{-\infty}^{+\infty} \frac{g(x)}{(u-x)} dx$$

It is well known (cf. [35]), that

$$\int_{-\infty}^{+\infty} |\hat{g}(u)|^2 du = \int_{-\infty}^{+\infty} |g(x)|^2 dx$$

If the function $g(x)$ is bandlimited and sampled at twice the bandwidth, then the functions $\hat{g}(u)$ and $g(x)$, both of which have the same bandwidth, can only be square integrable if the sample values are all finite. If the sampled function $g(x)$ is square integrable, then its first derivative is also square integrable and bounded. Thus, the Hilbert transform of the derivative of $\phi(\rho, \theta)$ is bounded so that the integral in Eq. (2-16) is convergent for all square-integrable-sampled pathlength functions.

Required Sampling Rates

It is well known from sampling theory [39] that the Fourier transform of a continuous function can be exactly determined from sampled values collected at a finite rate if the function is bandlimited and the samples are taken at a rate equal to twice the bandwidth. This rate is called the Nyquist rate. Since the refractive index field is spatially limited, it is not bandlimited and the index field cannot be exactly recovered using discrete data. For many spatially limited functions, however, the Fourier components above a certain frequency are much less than the components below that frequency;

thus, errors introduced by not including these terms can be negligible. This limiting frequency may be considered to be an effective bandwidth.

If the sampling rate is less than twice the bandwidth, aliasing will introduce errors in the resultant transform. Figure 2-5 shows the effect of aliasing. The frequency components beyond the cutoff

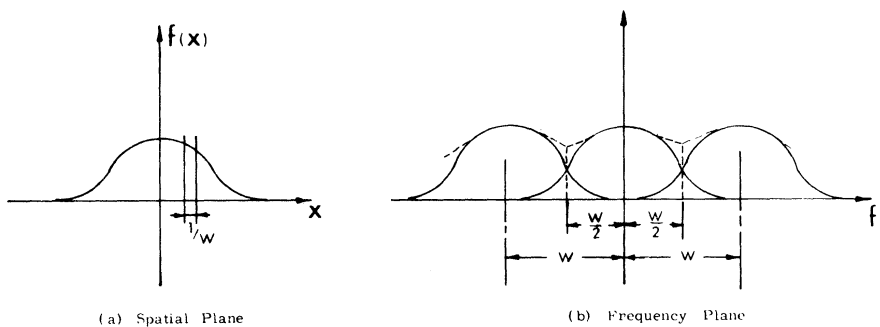


Fig. 2-5. Representation of the Aliasing Effects Caused by Undersampling. The Solid Lines in the Frequency Plane Represent the Actual Fourier Transform, the Dashed Lines Indicate the Resultant Transform Obtained by Undersampling

frequency impinge low frequencies. The transform obtained by undersampling will be significantly different from the actual transform if the frequency terms beyond the passband are comparable in magnitude to the components inside the passband.

If Eq. (2-10) is used to determine a line in the frequency plane of $f(x, y)$ from the sampled pathlength $\phi(\rho, \theta)$, then the separation between samples must be no greater than $(1/2B)$ where B is the effective bandwidth of the index field in the θ direction. Thus, the correct sampling rate is a function of the internal structure of the index field.

In addition to an adequate sampling rate in ρ , θ must be sampled at an adequate rate as well. Equations (2-17) to (2-20) are the basic relationships between a sampled two-dimensional function and the sampled two-dimensional transform. Of particular interest are the relationships between the series expansion of Eqs. (2-19) and (2-20). The spacing between sinc functions in the spatial plane, represented by Eq. (2-19), is $1/(2B_x)$ in the x direction and $1/(2B_y)$ in the y direction. Conversely, if the total extent of the refractive index field is L_x in the x direction and L_y in the y direction, then the spacing between sinc functions in the frequency plane, Eq. (2-20), is $1/L_x$ and $1/L_y$ in the f_x and f_y directions, respectively. The amplitudes of two adjacent sample points in the frequency plane are unrelated; between the sample points, however, the function varies smoothly due to the smooth behavior of the sinc function. Thus, the separation between sinc functions is a measure of the distance over which the series expansion in the frequency plane is approximately linear. The linear interpolation required to reformat the frequency plane will be acceptable if the separation between the two known points in the frequency plane is of the same order as the separation between the sinc functions of Eq. (2-20).

Figure 2-6 shows two radial lines in the frequency plane obtained from two adjacent views located near $\theta = 90^\circ$. The separation between sample points along each radial line is less than the separation between sinc functions if pathlength data has been collected at the

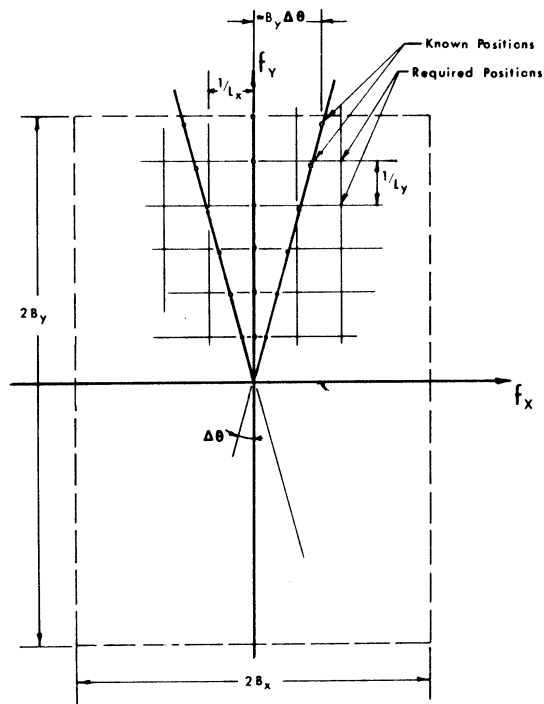


Fig. 2-6. Example of Interpolation Between Known Radial Lines in the Frequency Plane

Nyquist rate. The separation between the sample points on two adjacent radial lines increases with increasing frequencies. The separation between the radial lines at the frequency corresponding to the bandlimit determines the required sampling rate in θ . The condition to be satisfied is

$$B_y \tan(\Delta\theta) \leq \frac{1}{L_x} \quad (\text{for } \theta = \pi/2)$$

so that the allowable spacing between views is

$$\Delta\theta \leq \tan^{-1} \left(\frac{1}{\frac{B_y L_x}{y x}} \right) \approx \frac{1}{\frac{B_y L_x}{y x}} \quad (2-22)$$

Similar relationships can be used for other directions.

Consider a numerical example. Let the extent of the refractive index be 5 cm in the x direction and 5 cm in the y direction. Let the function be such that it can be accurately represented with sinusoidal functions of five cycles or less in the x directions; i.e., the bandwidth in the f direction is 1 cycle/cm ($B_x = 1$) and 10 cycles or less in the y direction; i.e., the bandwidth in the f_y direction is 2 cycles/cm ($B_y = 2$). Ten pathlength measurements in the $\theta = 0$ direction and 20 in the $\theta = 90^\circ$ direction are required. Near $\theta = 0$, the allowable spacing between views is approximately 12° ; and near the $\theta = 90^\circ$ direction, the allowable spacing is about 6° .

The required sampling rates have been discussed for the case where Eq. (2-10) is used to "build-up" the frequency plane. If the direct relationship between pathlength and the refractive index field, i.e., Eq. (2-16), is used, the sampling requirements remain the same. The convolution integral of Eq. (2-16) will include significant aliasing errors if the pathlength is not sampled at a rate equal to twice the effective bandwidth of the refractive index

field. The integral over θ cannot be accurately evaluated if the requirement for the allowable angular spacing between views, Eq. (2-22), is not satisfied. These requirements are independent of the numerical procedures used to evaluate the integrals (cf. [40]).

The maximum angular separation between consecutive views to allow linear interpolation in the frequency plane is determined by the separation between the two adjacent radial lines determined at the boundary of the frequency plane, as shown in Fig. 2-6. In regions near the origin of the frequency plane, the distance between the same adjacent lines is much less than the allowable separation so that, except at the highest frequency components, the frequency plane is oversampled. This appears to be wasteful in some sense and one is tempted to ask whether a more efficient scheme could be devised to reconstruct the index field with a smaller number of views. By analyzing the problem in terms of the space-bandwidth product, it can be shown that, in principle, the number of views required using Eq. (2-22) is at least twice the number actually required using more formal procedures.

The space-bandwidth product (SBWP) is a measure of the total number of degrees of freedom. For Eq. (2-19), for example, the space-bandwidth product is

$$\text{SBWP} = (2B_x L_x)(2B_y L_y) \quad (2-23)$$

which is equal to the total number of generating functions used in the expansion. Since the Fourier transform does not alter the number

of independent degrees of freedom, the SBWP of the series expansion for the frequency plane is also $(2B_y L_y)(2B_x L_x)$. (Note that the Fourier transform is complex so that there are two degrees of freedom for each sample point; but due to Hermitian symmetry, only half of the sample points are independent.)

If the total number of pathlength samples required by Eq. (2-22) to reconstruct the index field is calculated for a distribution with parameters L_x , L_y , B_x , and B_y , the result is, allowing for a few roundoff errors, $(8B_x B_y L_x L_y)$, which is equal to twice the SBWP of the original function. Thus, there are twice as many pathlength measurements required as degrees of freedom in the reconstruction. Each new independent pathlength measurement, however, provides new information about the refractive index field and should be capable of introducing a new degree of freedom into the reconstruction and, in principle, the number of ray samples required should be equal to the space-bandwidth product of the refractive index field. The formal procedure which can be used to realize the complete utilization of the data to reformat the frequency plane will be presented in the next chapter.

Since the refractive index field is considered to be bandlimited, the pathlength data collected for each direction of view will provide only $(2BL)$ independent pathlength measurements, where B is the bandwidth and L is the linear extent of the refractive index field normal to the direction of observation. For example, if the refractive index field can be expanded in terms of sinusoidal functions of 10 cycles and less in a direction normal to the direction of observation, then

there are only 20 independent pathlength measurements. The twenty-first measurement is a linear combination of the first 20 and represents redundant data.

Successive views of the index field must be taken until the total number of independent rays collected is equal to the space-bandwidth product. Since the number of independent rays varies from view to view, the total number of views required depends on the particular views selected. One collection of views that will provide only as many rays as the space-bandwidth product is to select every other view required by the interpolation scheme.

It is interesting to note that the view which is normal to the direction of the maximum value of (BL) has more degrees of freedom than views of other directions. In the next chapter, where the total angle of view available is less than 180° , the angle of view that is available will be centered about the direction with the greatest number of degrees of freedom.

Utilizing each independent ray to introduce a new degree of freedom in the restoration can only be realized if the system is free from errors. If errors are present, their effects can be greatly magnified in the reconstruction process. Since there are always errors present in a real system, the preceding arguments actually establish a theoretical lower limit on the number of pathlength measurements required. To minimize the effects of errors, redundant data must be collected. Since, in general, there are errors, the data will be inconsistent and the refractive index field must be

restored according to some criterion which will, hopefully,
smooth out the effects of noise.

CHAPTER 3

RECONSTRUCTION BY SERIES EXPANSION

The technique developed in the previous chapter can be utilized to reconstruct a three-dimensional refractive index field only if pathlength data can be collected over a full 180°. Unfortunately, a holographic system which allows a 180° angle of view will be very complex and requires numerous optical components. In addition, if an enclosed test cell is required, designing a cell which does not introduce severe optical distortions and avoids multiple reflections at the glass-air interface will be a formidable task. In this chapter, several techniques will be developed which do not require a 180° angle of view. These techniques will greatly simplify the eventual experimental application to temperature measurement.

Basic Approach

The procedures developed in this chapter all introduce a series expansion for $f(x, y)$ of the form

$$\bar{f}(x, y) = \sum_{m=1}^M \sum_{n=1}^N a_{mn} H_{mn}(x, y) \quad (3-1)$$

where $H_{mn}(x, y)$ are the generating functions and a_{mn} are their respective coefficients. Since only a finite number of terms are employed in the series expansion, the expansion will not, in general, converge to $f(x, y)$ exactly, but to an approximation of $f(x, y)$ which will be represented here as $\bar{f}(x, y)$.

When the series approximation of Eq. (3-1) is introduced into the basic integral equation, Eq. (2-3) becomes

$$\int_S \bar{f}(x, y) dS = \phi(\rho, \theta) \quad (3-2)$$

Combining Eq. (3-1) and Eq. (3-2) and interchanging the order of integration and summation produces the relationship

$$\sum_{m=1}^M \sum_{n=1}^N a_{mn} \int_S H_{mn}(x, y) dS = \phi(\rho, \theta) \quad (3-3)$$

The objective is to select the coefficients a_{mn} so that Eq. (3-3) is satisfied.

Equation (3-2) implies that

$$\int_S f(x, y) dS = \int_S \bar{f}(x, y) dS \quad (3-4)$$

Unless the refractive index field can be represented exactly with a finite number of terms in the series expansion of Eq. (3-1), the equality expressed by Eq. (3-4) will be satisfied only approximately. Thus, there is an inaccuracy introduced into Eq. (3-3).

There are doubtlessly a number of generating functions which can be employed and also a number of schemes to evaluate the coefficients a_{mn} . The complexity of Eq. (3-3) will be significantly reduced, however, if the generating functions $H_{mn}(x, y)$ can be easily integrated along any arbitrary straight line. A desirable choice for H_{mn} would

be a complete set of orthogonal functions. A set of orthogonal functions which can be properly integrated has been developed by Maldonado and Olsen [28] utilizing the schemes developed by Bhatia and Wolf [27]. This orthogonal series has been applied to interferometric measurements by Matulka and Collins [29]. To carry out the necessary orthogonalization, the pathlength data must be integrated over a full 180° .

Several series expansions and routines to evaluate the coefficients a_{mn} , will be developed and discussed here in detail; several other potential procedures will also be mentioned. The first technique analyzed, referred to as the grid method, represents the refractive index field as a set of rectangular elements.

Grid Method

A straightforward technique which can be applied to obtain a solution to Eq. (3-3) is to represent the refractive index field as a set of rectangular elements. Within each rectangular element, the refractive index has an unknown, but constant, value. As each ray passes through the refractive index field, various segments of the ray lie in each element. The length of each segment can be determined from geometrical considerations alone. If a ray segment of length l_i is in a cell with refractive index n_i , the pathlength of the ray within the rectangular element is $n_i l_i$. As the ray passes through other elements of the grid, there are similar contributions to the pathlength. The total pathlength for a ray is the sum of these contributions. Thus, the total pathlength of the ray can be expressed as a linear equation; the refractive index of each element is the

dependent variables. By considering a number of independent rays equal to the number of grid elements, a linear system of algebraic equations can be generated and solved for the unknown refractive index values.

More insight into the grid method can be obtained, however, if the technique is derived in a more formal manner. The series representation for the refractive index field is

$$\bar{f}(x, y) = \sum_{m=1}^M \sum_{n=1}^N \bar{f}(\ell_x m, \ell_y n) \cdot \text{rect}\left[\frac{1}{\ell_x}(x - \ell_x m)\right] \text{rect}\left[\frac{1}{\ell_y}(y - \ell_y n)\right] \quad (3-5)$$

where

$$\text{rect}(x) \equiv \begin{cases} 1 & \text{if } |x| \leq 0.5 \\ 0 & \text{otherwise} \end{cases}$$

and ℓ_x and ℓ_y are shown in Fig. 3-1. Consider a shifted coordinate axis defined as (Fig. 3-1)

$$x' = x - m\ell_x$$

$$y' = y - n\ell_y$$

The equation of an arbitrary ray with slope a_i and y intercept b_i is

$$y = a_i x + b_i \quad (3-6)$$

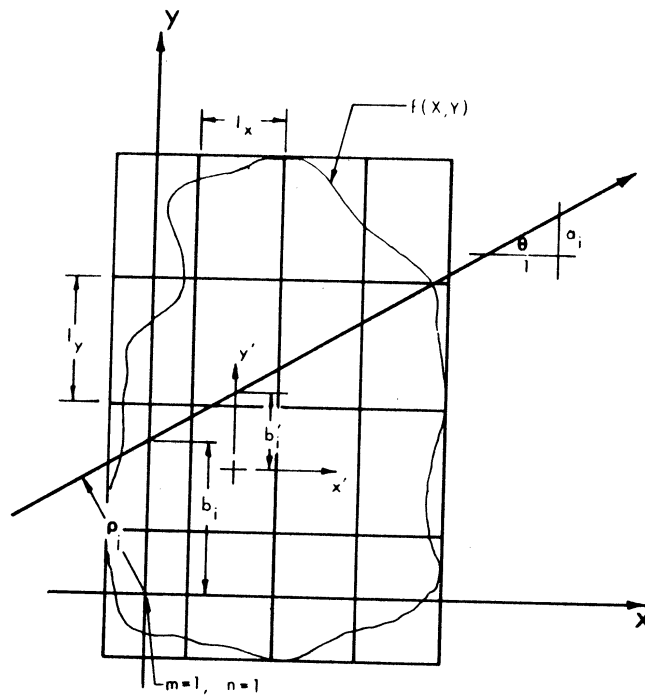


Fig. 3-1. Nomenclature for the Grid Series Expansion of the Refractive Index Field

or, in the shifted coordinate system,

$$y' = a_i x' + b'_i$$

The differential pathlength dS is

$$dS = (1 + a_i^2)^{1/2} dx$$

Equation (3-3) becomes, extending the limits of integration to infinity,

$$\sum_{m=1}^M \sum_{n=1}^N \bar{f}(\ell_x m, \ell_y n) \int_{-\infty}^{+\infty} \text{rect}\left(\frac{1}{\ell_x} x'\right) \text{rect}\left[\frac{1}{\ell_y} (a_i x' + b'_i)\right] \left(1 + a_i^2\right)^{1/2} dx' = \phi(\rho, \theta) \quad (3-7)$$

By introducing the dummy variable $z = -b'_i/a_i$ and using the evenness of the rect function, Eq. (3-7) can be written as

$$\left(1 + a_i^2\right)^{1/2} \sum_{m=1}^M \sum_{n=1}^N \bar{f}(\ell_x m, \ell_y n) \int_{-\infty}^{+\infty} \text{rect}\left(\frac{1}{\ell_x} x'\right) \text{rect}\left[\frac{a_i}{\ell_y} (z - x')\right] dx' = \phi(\rho, \theta) \quad (3-8)$$

The integral in Eq. (3-8) is a convolution integral which is easily evaluated graphically. The cases for $a_i = 0$ and $|a_i| = \infty$ have to be handled as special cases. For summarizing the results for all values of $|a_i|$, it is convenient to note the following geometric relations

$$b'_i = b_i + a_i m \ell_x - n \ell_y = \rho_i \sec \theta_i + a_i m \ell_x - n \ell_y$$

$$c'_i = c_i - m \ell_x = \rho_i + m \ell_x$$

where $x' = c'_i$ is the equation of a ray for which $|a_i| = \infty$. The algebraic equations which can be used to determine $\bar{f}(\ell_x m, \ell_y n)$ are

$$\sum_{m=1}^M \sum_{n=1}^N W(a_i, \rho_i) \bar{f}(\ell_x m, \ell_y n) = \phi(\rho, \theta) \quad (3-9)$$

where

$$W(a_i, \rho_i) = \left\{ \begin{array}{ll} (1 + a_i^2)^{1/2} \ell_x & \text{for } |b'_i| \leq \frac{\ell_y - \ell_x |a_i|}{2} \quad \text{and} \\ & |a_i| \leq \frac{\ell_y}{\ell_x} \\ (1 + a_i^2)^{1/2} \frac{\ell_y}{|a_i|} & \text{for } |b'_i| \leq \frac{\ell_x |a_i| - \ell_y}{2} \quad \text{and} \\ & |a_i| > \frac{\ell_y}{\ell_x} \\ \frac{(1 + a_i^2)^{1/2}}{|a_i|} \left(\frac{\ell_x |a_i| + \ell_y}{2} - |b'_i| \right) & \text{for} \\ & \frac{|\ell_y - \ell_x |a_i||}{2} < |b'_i| \leq \frac{|\ell_y + \ell_x a_i|}{2} \\ \ell_x & \text{for } |c'_i| < \frac{\ell_x}{2} \quad \text{and } |a_i| = \infty \\ 0 & \text{for } |b'_i| > \frac{\ell_x |a_i| + \ell_y}{2} \end{array} \right. \quad (3-10)$$

There is no requirement to collect rays over a complete 180° angle of view in order to solve the linear system of equations. It is particularly important to note that Eq. (3-10) reveals a

fundamental result that applies to all reconstruction techniques; the actual angle of view available is a relative factor. The coefficients in Eq. (3-10) are determined by the ratio of a_i to ℓ_y/ℓ_x . Consider an index field which is expanded in a series with M elements in the x direction and N elements in the y direction; if data is collected from views centered about the $\theta = 0$ direction, then an "effective" angle of view can be defined as

$$\theta_{\text{eff}} = \tan^{-1} \left(a_i \frac{L_x}{M} \frac{N}{L_y} \right) = \tan^{-1} \left(\tan \theta \frac{L_x}{M} \frac{N}{L_y} \right) \quad (3-11)$$

The term $\left(\frac{L_x}{M} \frac{N}{L_y} \right)$ can be considered to be an obliquity factor. The actual angle of view and the effective angle of view are only equal if $\ell_x = \ell_y$ (see Fig. 3-1). Two experiments with the same effective angle of view can produce the same matrix W_{ij} even though the actual angles of view may be quite different. The effects of input errors on the errors in the reconstruction are determined by the characteristics of the matrix W_{ij} ; thus, the effective angle of view is actually a more relevant measurement than the actual angle of view.

A number of independent rays equal to the number of grid elements can be collected and solved for $\bar{f}(x, y)$. This is in agreement with the argument that the number of rays required is equal, in principle, to the space-bandwidth product of the refractive index field which, in this case, is equal to the number of grid elements. This procedure was modeled numerically and it was found that the resulting errors are a strong function of the grid spacing used and the particular rays chosen. In general, the errors were unacceptably large.

Fortunately, collecting many more pathlength samples than unknowns, i.e., collecting redundant data, will greatly improve the accuracy of the grid method. Implementing redundant data will be discussed in detail later in this chapter.

Sample Method

Dividing the refractive index field into a set of quantized elements yields a crude representation of a function which is known to be smoothly varying and continuous. The integral of the quantized representation and the integral of the actual function may be considerably different; i.e., the approximation of Eq. (3-4) is unsatisfactory. In addition, the grid expansion does not distinguish between points within the same grid element; it does not, by itself, provide an interpolation procedure to determine values located, for example, near the boundary of two grid elements. Dividing the refractive index field into a greater number of elements is an unacceptable solution since, as will be shown, the resulting equation set becomes poorly conditioned and the amount of input data available in most interferometric studies does not warrant a higher order expansion.

The technique presented here, and referred to as the sample method, alleviates the basic cause of these problems by using a more natural expansion for $f(x, y)$. The expansion used is based on the Whittaker-Shannon sampling theorem [39]. This theorem states that if a function is bandlimited, it can be exactly represented by

$$g(x, y) = \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} g\left(\frac{m}{2B_x}, \frac{n}{2B_y}\right) \cdot \operatorname{sinc}\left[2B_x\left(x - \frac{m}{2B_x}\right)\right] \operatorname{sinc}\left[2B_y\left(y - \frac{n}{2B_y}\right)\right] \quad (3-12)$$

where B_x and B_y are the bandwidths of the function in the x and y direction, respectively, $g\left(\frac{m}{2B_x}, \frac{n}{2B_y}\right)$ are the values of $g(x, y)$ at the sample points, and $\operatorname{sinc}(x) \equiv \left[\sin(\pi x)\right]/\pi x$. In the method proposed here, a truncated form of Eq. (3-12) is used to represent $f(x, y)$. Since $f(x, y)$ is spatially bounded, it cannot actually be strictly bandlimited. The series of Eq. (3-12) can still be used as an approximation to $f(x, y)$, however. The approximation $\bar{f}(x, y)$ can be represented as

$$\bar{f}(x, y) = \sum_{m=1}^M \sum_{n=1}^N \bar{f}(\ell_x m, \ell_y n) \cdot \operatorname{sinc}\left[\frac{1}{\ell_x}(x - \ell_x m)\right] \operatorname{sinc}\left[\frac{1}{\ell_y}(y - \ell_y n)\right] \quad (3-13)$$

where ℓ_x and ℓ_y are shown in Fig. 3-2. Equation (3-3) becomes

$$\sum_{m=1}^M \sum_{n=1}^N \bar{f}(\ell_x m, \ell_y n) \int_{S_i} \operatorname{sinc}\left[\frac{1}{\ell_x}(x - \ell_x m)\right] \cdot \operatorname{sinc}\left[\frac{1}{\ell_y}(y - \ell_y n)\right] dS_i = \phi(\rho, \theta) \quad (3-14)$$

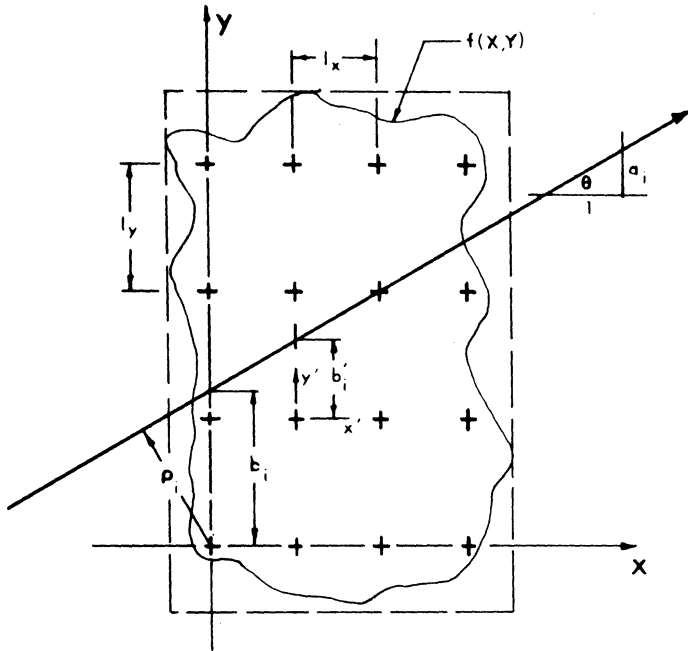


Fig. 3-2. Nomenclature for the Sample Series Expansion of the Refractive Index Field

Equation (3-14) can be written as

$$\sum_{m=1}^M \sum_{n=1}^N \bar{f}(l_x^m, l_y^n) \int_{-\infty}^{+\infty} \text{sinc}\left(\frac{x - l_x^m}{l_x}\right) \cdot$$

$$\text{sinc}\left(\frac{y - l_y^n}{l_y}\right) \left(1 + a_i^2\right)^{1/2} dx = \phi(\rho, \theta) \quad (3-15)$$

where the limits of integration have been extended to $\pm\infty$. For convenience, a coordinate system,

$$x' = x - \ell_x m$$

$$y' = y - \ell_y n$$

centered on each sampling point is defined. The equation of any ray can then be written as

$$x' = a_i x' + b_i'$$

(see Fig. 3-2), and Eq. (3-15) becomes

$$\sum_{m=1}^M \sum_{n=1}^N \bar{f}(\ell_x m, \ell_y n) \int_{-\infty}^{+\infty} \text{sinc} \left(\frac{x'}{\ell_x} \right) \text{sinc} \left(\frac{a_i x' + b_i'}{\ell_y} \right) \left(1 + a_i^2 \right)^{1/2} dx' = \phi(\rho, \theta) \quad (3-16)$$

If $a_i \neq 0$, Eq. (3-16) can be rewritten as

$$\left(1 + a_i^2 \right)^{1/2} \sum_{m=1}^M \sum_{n=1}^N \bar{f}(\ell_x m, \ell_y n) \int_{-\infty}^{+\infty} \text{sinc} \left(\frac{x'}{\ell_x} \right) \cdot \text{sinc} \left[\frac{a_i}{\ell_x} (q_i' - x') \right] dx' = \phi(\rho, \theta) \quad (3-17)$$

where $q_i' = -b_i'/a_i$ and where the symmetry of the sinc function has been used. The integral in Eq. (3-17) is a convolution; it can

readily be evaluated by use of elementary properties of the two-dimensional Fourier transform

$$\int_{-\infty}^{+\infty} \text{sinc} \left(\frac{x'}{\ell_x} \right) \text{sinc} \left[\frac{a_i}{\ell_y} (q'_i - x') \right] dx' \\ = \mathcal{F}^{-1} \left[\frac{|\ell_x| |\ell_y|}{|a_i|} \text{rect} \left(\ell_x f_q \right) \text{rect} \left(\frac{\ell_y}{a_i} f_q \right) \right] \quad (3-18)$$

where

$$\text{rect} (x) \equiv \begin{cases} 1 & \text{if } |x| \leq 1/2 \\ 0 & \text{if } |x| > 1/2 \end{cases}$$

is the Fourier transform of $\text{sinc}(x)$. Evaluating the right-hand side of Eq. (3-18) and substituting it, along with $q'_i = -b'_i/a_i$, into Eq. (3-17) results in

$$\left(1 + a_i^2\right)^{1/2} \sum_{m=1}^M \sum_{n=1}^N \bar{f}(\ell_x m, \ell_y n) \ell_x \text{sinc} \left(\frac{b'_i}{\ell_x} \right) = \phi(\rho, \theta) \quad (3-19)$$

if

$$0 < \left(\frac{|a_i|}{\ell_y} \right) \leq \frac{1}{\ell_x}$$

or

$$\left(1 + a_i^2\right)^{1/2} \sum_{m=1}^M \sum_{n=1}^N \bar{f}(\ell_x m, \ell_y n) \frac{\ell_y}{|a_i|} \text{sinc} \left(\frac{b'_i}{a_i \ell_x} \right) = \phi(\rho, \theta) \quad (3-20)$$

if

$$\frac{1}{l_x} < \frac{|a_i|}{l_y}$$

Rays for which $|a_i| = 0$ or $|a_i| = \infty$ must be handled as special cases.

For summarizing the results for all values of $|a_i|$, it is convenient to introduce the following geometric relations:

$$b'_i = b_i + a_i m l_x - n l_y = \rho_i \sec \theta_i + a_i m l_x - n l_y \quad (3-21)$$

$$c'_i = c_i - m l_x = \rho_i + m l_x$$

where $x' = c'_i$ is the equation of a ray for which $|a_i| = \infty$. The algebraic equations which can be used to determine $\bar{f}(l_x m, l_y n)$ are

$$\sum_{m=1}^M \sum_{n=1}^N W(a_i, \rho_i) \bar{f}(l_x m, l_y n) = \phi(\rho, \theta) \quad (3-22)$$

where

$$W(a_i, \rho_i) = \begin{cases} \left(1 + a_i^2\right)^{1/2} |l_x| \operatorname{sinc}\left(\frac{\rho_i \sec \theta_i + a_i l_x m - l_y n}{l_y}\right) & \text{for} \\ & 0 \leq |a_i| \leq \frac{|l_y|}{|l_x|} \\ \left(1 + a_i^2\right)^{1/2} \frac{|l_y|}{|a_i|} \operatorname{sinc}\left(\frac{\rho_i \sec \theta_i + a_i l_x m - l_y n}{a_i l_x}\right) & \text{for} \\ & \frac{|l_y|}{|l_x|} < |a_i| < \infty \\ \operatorname{sinc}\left(\frac{\rho_i + m l_x}{l_x}\right) & \text{for } |a_i| = \infty \end{cases}$$

Equation (3-22) provides the basis of a scheme for analyzing multidirectional interferometric data. As shown in Fig. 3-2, the region of a test section in which significant refractive-index variation occurs is considered to lie within a rectangular region containing $(M \times N)$ sample points. If the differential optical pathlength, ϕ , of $(M \times N)$ different rays traversing this region is measured, Eq. (3-22) yields a set of $(M \times N)$ linear algebraic equations which can be solved for the $(M \times N)$ unknowns, $\bar{f}(\ell_x^m, \ell_y^n)$. These values can then be substituted into Eq. (3-13) to obtain an approximate representation of $f(x, y)$ anywhere in the particular plane being studied.

Once again the elements of the matrix W_{ij} are functions of the ratio of $|a_i|$ to ℓ_y/ℓ_x , and an obliquity factor can be defined to determine an effective angle of view which is identical to Eq. (3-11).

Frequency Plane Restoration

The following technique, referred to as the frequency plane restoration technique, is based on the concepts developed by J. L. Harris [15] for superresolution. The possibility of using a reconstruction procedure based on Harris's work was mentioned but not developed by Hoppe [16] in the context of electron microscopy. Harris's technique constitutes, in theory, a procedure for obtaining higher frequency components and, hence, higher resolution than is apparently available from an aperture of limited extent. In practice, these techniques are subject to severe limitations imposed by noise. In this application, the effects of system noise will be minimized by using redundant data and truncating the resultant series expansion.

Suppose that the pathlengths are known for values of θ between the limits $\pm\beta$, as shown in Fig. 3-3a. Using the techniques developed in Chapter 2, points in the frequency plane are determined along radial lines over an angular range of 2β , as shown in Fig. 3-3b. The objective is to determine the frequency components outside of this known region. Such a procedure is possible because the Fourier

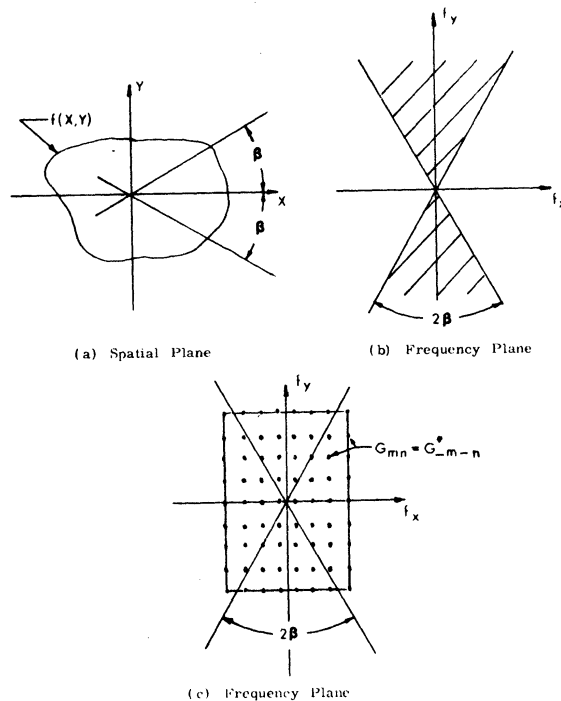


Fig. 3-3. Nomenclature for Frequency Plane Restoration Technique

transform of the spatially limited function $f(x, y)$ is analytic. If the Fourier transform can be uniquely determined over any finite domain, then the entire frequency plane can be determined using analytic continuation. The procedure employed here is not a classical analytic continuation which uses a Taylor series expansion, but is an equivalent technique which utilizes the sampling theorem in the frequency plane.

Let $f(x, y)$ be bounded in the x direction by $\pm L_x/2$ and in the y direction by $\pm L_y/2$; then, $f(x, y)$ can be described completely in the bounded region by the two-dimensional Fourier series

$$f(x, y) = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} G_{mn} e^{+i2\pi\left[\left(\frac{mx}{L_x}\right) + \left(\frac{ny}{L_y}\right)\right]} \quad (3-23)$$

where

$$G_{mn} = \frac{1}{L_x L_y} \int_{-L_y/2}^{+L_y/2} \int_{-L_x/2}^{+L_x/2} f(x, y) e^{-i2\pi\left[\left(\frac{mx}{L_x}\right) + \left(\frac{ny}{L_y}\right)\right]} dx dy \quad (3-24)$$

The relationship between the Fourier spectrum of $f(x, y)$ and the coefficients G_{mn} can be found by substituting Eq. (3-23) into Eq. (2-6). The result, after some manipulation, is

$$F(f_x, f_y) = L_x L_y \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} G_{mn} \operatorname{sinc}\left[L_x\left(\frac{m}{L_x} - f_x\right)\right] \operatorname{sinc}\left[L_y\left(\frac{n}{L_y} - f_y\right)\right] \quad (3-25)$$

Thus, the coefficients G_{mn} are equal to the sampled values of the Fourier transform at the locations $f_x = m/L_x$ and $f_y = n/L_y$; the positions of the G_{mn} coefficients are shown in Fig. 3-3c.

If the summation in Eq. (3-25) is truncated, Eq. (3-25) can be used to generate a set of simultaneous algebraic equations of the form

$$W_{ij} G_j = F_i \quad (3-26)$$

where the column vector F_i consists of the sampled values of the frequency plane along the known radial lines. The equation set can be solved for the column vector G_j whose elements are the unknown G_{mn} coefficients. Since the G_{mn} coefficients can be determined for locations which do not lie within the domain covered by the known radial lines, the two-dimensional Fourier transform $F(f_x, f_y)$ can be determined and $f(x, y)$ can be obtained by evaluating Eq. (3-23).

If there are to be $(M \times N)$ degrees of freedom in the reconstructed refractive index field, there are $(M \times N)$ coefficients which must be determined in the frequency plane. Since the coefficients G_{mn} are complex, there are $2(M \times N)$ discrete numbers which must be specified. Due to Hermitian symmetry, not all of the coefficients are independent, however. The Hermitian symmetry can be expressed as

$$G_{mn} = G_{-m-n}^*$$

where the * denotes complex conjugate. Since Eq. (3-26) is linear, the real and imaginary parts of the G_{mn} coefficients may be expanded

in separate sets of linear equations. Thus, it is necessary to solve two systems of equations, each with $(M \times N)/2$ unknowns.

From a computational viewpoint, this technique has two advantages over the grid and sample methods. The number of computer operations necessary to solve a system of linear equations with K unknowns is proportional to K^3 . Thus, solving two systems of equations with $(K/2)$ unknowns requires four-fold fewer operations than solving one equation set with K unknowns. In addition, the computer memory required to store a matrix with $(K/2)$ rows and $(K/2)$ columns is one-fourth that required by a K -row and K -column matrix.

In practice, any error in the known values of $F(f_x, f_y)$ can be greatly magnified in the resultant G_{mn} . The procedure can be modified to reduce the effects of errors by truncating the summation in Eq. (3-26) to include only the coefficients within a limited rectangular region about the origin and by using redundant data. The next section summarizes the implementation and importance of redundant data, not only for this procedure, but for all of the reconstruction techniques.

Use of Redundant Data

In theory, the number of degrees of freedom in the reconstructed refractive index field can be equal to the number of independent pathlength samples. In practice, however, any errors in the pathlength data can be greatly amplified by the reconstruction process.

Since the refractive index field is expanded in a finite number of terms, the path integral of the truncated series expansion and

the integral of the actual index field will not be equal in general. The difference between these two integrals can be considered to be an error in the pathlength data. Thus, there is an error that results from the analysis technique itself. The same error also exists in the Fourier synthesis techniques developed in the previous chapter. In the Fourier techniques, the error arises from the discrete sampling rate of the pathlength data which introduces aliasing effects. Using the grid method, the sample methods, or the frequency plane restoration technique, values of M and N must be selected so that the series expansion is capable of approximating the actual refractive index field.

The use of redundant data can greatly reduce the effects of this systematic error by averaging several samples. If the errors for each pathlength sample are uncorrelated, averaging will suppress the error and retain only the desired output signal. Thus, redundancy is a requirement for good reconstructions. The amount of redundancy that is needed depends on the characteristics of the reconstruction process. In general, greater redundancy is required as the angle of view decreases. This will be discussed more fully in the next section.

Redundancy is introduced by collecting many more pathlength samples than degrees of freedom in the series expansion. The result is a system of overdetermined equations

$$W_{ij} f_j = \phi_i \quad (3-27)$$

The number of rows in the matrix W_{ij} is equal to the number of pathlength samples and the number of columns is equal to the number of terms used in the series expansion. Since there are errors in the pathlength samples, the system of equations will be inconsistent. The overdetermined system, therefore, has no exact solution. A "solution" which is often used, however, is the solution vector f_j for which

$$||W_{ij}f_j - \phi_i|| = \min \quad (3-28a)$$

and

$$||f_j|| = \min \quad (3-28b)$$

where $||\dots||$ denotes the Euclidean norm defined as

$$||x_i|| = (x_i^t x_i)^{1/2} = \left[\sum_{i=1}^N (x_i)^2 \right]^{1/2}$$

and ϕ_i is a column vector consisting of the pathlength data.

Equations (3-28a) and (3-28b) define the least-mean-square solution.

If W_{ij} is full rank, then Eq. (3-28a) is sufficient to uniquely specify f_j and Eq. (3-28b) is not required.

The least-mean-square solution to an overdetermined system of algebraic equations is (cf. [41]),

$$f_j = (W_{ji} W_{ij})^{-1} W_{ji} \phi_i \quad (3-29)$$

Although Eq. (3-29) can be used directly to obtain the least-mean-square solution, it is not the most numerically accurate procedure.

Evaluating the inner product $(W_{ji} W_{ij})$ can introduce unnecessary numerical inaccuracies which can be particularly large when the matrix W_{ij} is poorly conditioned. The errors introduced by evaluating the inner product can be eliminated by using a procedure which works directly with the system of overdetermined equations and does not explicitly form the inner product. Such a procedure has been developed by Golub [42]. Golub's method is a stable numerical procedure utilizing orthogonal Householder reflections. Since only orthogonal transformations are applied to W_{ij} , there are no large round-off errors accumulated. (An Algol Program for this procedure is given in Ref. 43, page 113.)

An indication of the amount of redundant data that is required can be obtained by considering the effects of redundancy on the convergence of the various series solutions. Consider the frequency plane restoration technique. If the Fourier series coefficients G_{mn} are determined for positions within a rectangular region centered about the origin of the frequency domain, then the inverse transform of the sampling series generated by the G_{mn} coefficients will be a bandlimited representation of $f(x, y)$. In order to obtain a true bandlimited reconstruction of $f(x, y)$, the one-dimensional Fourier transform of the pathlength data must be free from aliasing effects up to the frequency corresponding to the frequency limit used in the series expansion. If the refractive index field has an effective bandlimit of B and the bandlimit of the series expansion is B_ℓ , then the pathlength data must be collected at a rate of $(B + B_\ell)/2$. This will require collecting more pathlength samples than degrees of

freedom in the reconstructed index field; i.e., redundant data is required. Sampling at this rate does not eliminate errors, however, since the bandlimit of the refractive index field is actually infinite and, consequently, there are always some aliasing effects.

The sample method also converges to a bandlimited version of the input function if pathlength data is collected at an average rate of $(B_x + B_y)/2$. Brown [44] has shown that given an arbitrary function $f(x, y)$, the bandlimited function $\bar{f}(x, y)$ which satisfies

$$\iint_{-\infty}^{+\infty} |f(x, y) - \bar{f}(x, y)|^2 dx dy = \min \quad (3-30)$$

is the bandlimited representation of $f(x, y)$, i.e.,

$$\bar{f}(x, y) = f(x, y) * \left[\text{sinc}\left(\frac{x}{2B_x}\right) \text{sinc}\left(\frac{y}{2B_y}\right) \right] \quad (3-31)$$

where B_x and B_y are the bandwidths of $\bar{f}(x, y)$ and the $*$ denotes the convolution. The sample method actually determines an $\bar{f}(x, y)$ which satisfies the following relation

$$\sum_{i=1}^I \left\{ \int_{-\infty}^{+\infty} [f(x, y) - \bar{f}(x, y)] dS_i \right\}^2 = \min \quad (3-32)$$

where

$$\int_{-\infty}^{+\infty} \bar{f}(x, y) dS_i = W_{ij} f_j$$

The area integral of Eq. (3-30) is replaced by a summation over a collection of line integrals. It is reasonable to expect that if

Eq. (3-32) is satisfied for a sufficiently large number of line integrals, then Eq. (3-30) is satisfied also and vice-versa. Thus, the sample method also converges to a bandlimited representation of $f(x, y)$. As was seen in the discussion of the convergence of the frequency plane restoration technique, the pathlength samples do not contain the information necessary to reconstruct a bandlimited representation of the index field unless the samples are taken at a rate of $(B_\ell + B)/2$.

The convergence of the grid method is not so apparent. Since the overdetermined equation set is solved for the least-mean-square solution, it is reasonable to expect that if a sufficient number of pathlength samples are taken, the coefficient of each grid element will converge to the root-mean-square value within the grid.

The sampling rates required for convergence of the series solutions actually represent the minimum allowable redundant data that can be used. The linear transformation described by the matrix W_{ij} may amplify the input errors. Least-mean-square solutions only average over several signals; as will be shown, effects of uncorrelated errors are suppressed at a rate proportional to the square root of the number of redundant pathlength measurements. If the errors are correlated, then the effectiveness of the least-mean-square solution will be decreased in proportion to the degree of correlation of the errors. The next section compares the effect of errors in the grid and sample methods; a similar comparison could be obtained for the frequency plane restoration technique.

Comparison of Grid and Sample Methods

Several numerical parameters may be determined which indicate the basic effect of input errors on the sample and grid methods. The various parameters differ in the specific assumptions made about the characteristics of the errors and the structure of the refractive index field. It would be desirable to have an indication of the effect of the errors which is independent of the input function and the statistical nature of the errors. Although such a parameter is available, called the condition, it will be an indication of the greatest possible effect of errors on the reconstruction. To obtain a more realistic indication of the effects of errors, certain reasonable assumptions will be made about the errors in the pathlength samples.

It is not possible to obtain a comparison of the techniques which clearly reveals one technique to be superior to the other. The success of either technique is a function of the particular index field to which it is applied. In this sense, the most comprehensive comparisons can be obtained by selecting a temperature field which is similar in structure to the temperature fields likely to be encountered experimentally and simulating the reconstruction process; this is the subject of the next chapter. The comparisons that follow are of value because they reveal the increased effects of errors as either the number of degrees of freedom (resolution) is increased or the angle of view is decreased.

The sample and grid method both require solving an overdetermined and inconsistent set of equations

$$\overline{W}\overline{f} = \overline{\phi} \quad (3-33)$$

where the subscript notation has been replaced by the vector notation. \overline{W} is defined by Eq. (3-9) for the grid method and Eq. (3-22) for the sample method. The overdetermined equation set generated can be considered to be a linear system, as shown in Fig. 3-4. A common measure

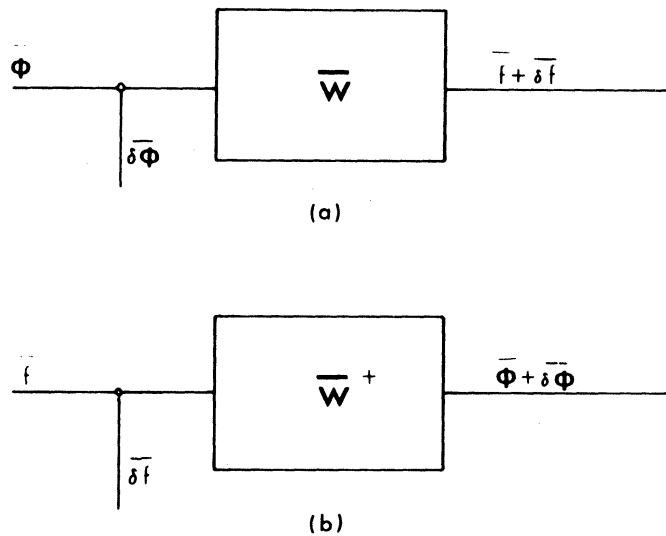


Fig. 3-4. Linear System Representation of the Transformation $\overline{W}\overline{f} = \overline{\phi}$

of the significance of errors, referred to as noise in a linear systems analysis, is the signal-to-noise ratio where the signal is either the pathlength samples $\bar{\phi}$ or the series coefficients \bar{f} . The signal-to-noise ratio is defined here as the ratio of the root-mean-square energy of the signal to the root-mean-square energy of the noise. The signal-to-noise ratio can be expressed as

$$(S/N)_{\bar{f}} = \frac{||\bar{f}||}{||\bar{\delta f}||} \quad (3-34)$$

$$(S/N)_{\bar{\phi}} = \frac{||\bar{\phi}||}{||\bar{\delta\phi}||}$$

where $\bar{\delta f}$ and $\bar{\delta\phi}$ represent the noise present in the series coefficients and the pathlength samples, respectively, as indicated in Fig. 3-4.

A quantity of interest is the ratio of $(S/N)_{\bar{\phi}}$ to $(S/N)_{\bar{f}}$. This ratio is an indication of the increase or decrease of the noise level in the reconstruction. For example, if the ratio is less than one, then the relative uncertainty in the reconstruction is less than the relative uncertainty in the pathlength samples. This ratio is called, for reference, the noise amplification factor, R.

It is necessary to introduce the following relations:

$$\begin{aligned} ||\bar{f}|| &= (\bar{f}^t \bar{f})^{1/2} \\ ||\bar{\delta f}|| &= (\bar{\delta f}^t \bar{\delta f})^{1/2} \\ ||\bar{\phi}|| &= (\bar{\phi}^t \bar{\phi})^{1/2} = [(\bar{W}\bar{f})^t (\bar{W}\bar{f})]^{1/2} \leq ||\bar{W}|| ||\bar{f}|| \\ ||\bar{\delta\phi}|| &= (\bar{\delta\phi}^t \bar{\delta\phi})^{1/2} = [(\bar{W}\bar{\delta f})^{1/2} (\bar{W}\bar{\delta f})]^{1/2} \leq ||\bar{W}|| ||\bar{\delta f}|| \end{aligned} \quad (3-35)$$

where the superscript t denotes the transpose and the inequalities in the last two equations have been introduced by applying Schwarz's inequality. The ratio $(S/N)_{\phi}^{-}/(S/N)_{R}^{-}$ is

$$R = \frac{(S/N)_{\phi}^{-}}{(S/N)_{\bar{f}}^{-}} = \frac{||\bar{W}\bar{f}||/||\bar{W}\bar{\delta f}||}{||\bar{f}||/||\bar{\delta f}||} = \frac{||\bar{W}\bar{f}|| \quad ||\bar{\delta f}||}{||\bar{f}|| \quad ||\bar{W}\bar{\delta f}||} \quad (3-36)$$

Although Eq. (3-36) represents the general result for R , evaluating Eq. (3-36) requires a knowledge of \bar{f} and $\bar{\delta f}$ which are not, in general, available. It is possible, however, to obtain values for R which involve making certain assumptions about \bar{f} and/or $\bar{\delta f}$; it is also possible to obtain an upper bound on R which is independent of \bar{f} and $\bar{\delta f}$.

It is well known (cf. [45]) that a real matrix \bar{W} with I rows and J columns ($I \geq J$) can be decomposed into the form

$$\bar{W} = \bar{U}\bar{\Sigma}\bar{V}^t \quad (3-37)$$

where

\bar{U} = a matrix consisting of the orthonormalized
(i.e., $\bar{U}^t\bar{U} = I$) eigenvectors associated with $(W\bar{W}^t)$

$\bar{\Sigma}$ = a diagonal matrix of the singular values of \bar{W} ,
 $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_j)$

\bar{V} = a matrix consisting of the orthonormalized eigenvectors
of W^tW

The least-mean-square solution of an equation set $\bar{W}\bar{f} = \bar{\phi}$ is

$$\bar{f} = \bar{W}^+\bar{\phi} \quad (3-38)$$

The superscript + denotes the pseudoinverse

$$W^+ = (\overline{U\Sigma V^t})^+ = V\Sigma^+U^t \quad (3-39)$$

where

$$\Sigma^+ = \left[\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_j) \right]^+ = \text{diag}\left(\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots, \frac{1}{\lambda_j}\right)$$

Equation (3-38) becomes

$$\overline{f} = \overline{V}\Sigma^+\overline{U}^t\overline{\phi} \quad (3-40)$$

and since the system is linear,

$$\overline{\delta f} = \overline{V}\Sigma^+\overline{U}^t\overline{\delta\phi} \quad (3-41)$$

Evaluating the norm of $\overline{\delta f}$

$$\begin{aligned} (\overline{\delta f}^t\overline{\delta f})^{1/2} &= \left(\overline{\delta\phi}^t\overline{U}\Sigma^+\overline{V}^t\overline{V}\Sigma^+\overline{U}^t\overline{\delta\phi} \right)^{1/2} \\ &= \left(\overline{\delta\phi}^t\overline{U}\overline{\Lambda}\overline{U}^t\overline{\delta\phi} \right)^{1/2} \end{aligned} \quad (3-42)$$

where

$$\overline{\Lambda} = \text{diag}\left(\frac{1}{\lambda_1^2}, \frac{1}{\lambda_2^2}, \dots, \frac{1}{\lambda_j^2}\right)$$

the λ_i^2 are the eigenvalues associated with W^tW .

In the absence of any other information, a reasonable assumption for the elements of $\overline{\delta\phi}$ is that they are uncorrelated random variables with identical distributions; that is,

$$E(\delta\phi_j \delta\phi_k) = \begin{cases} \sigma^2 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases} \quad (3-43)$$

where E denotes the expected value. Combining Eqs. (3-43) and (3-42) and using the orthonormal properties of \bar{U} , results in

$$||\bar{\delta f}|| = (\bar{\delta f}^t \bar{\delta f})^{1/2} = \left[\sum_{j=1}^J \left(\frac{1}{\lambda_j} \right) \sigma^2 \right]^{1/2} \quad (3-44)$$

Equation (3-36) then becomes

$$R = \frac{||\bar{W}\bar{f}||}{||\bar{f}||} \frac{\left[\sum_{j=1}^J \left(\frac{1}{\lambda_j} \right) \right]^{1/2}}{(I)^{1/2}} \quad (3-45)$$

Equation (3-45) shows that the effect of noise is decreased in proportion to the square root of the number of pathlength samples.

If Schwarz's inequality is used in Eq. (3-45), then

$$R \leq \frac{||W||}{(I)^{1/2}} \left[\sum_{j=1}^J \left(\frac{1}{\lambda_j} \right) \right]^{1/2} = \frac{\left[\sum_{j=1}^J \left(\frac{\lambda_{\max}}{\lambda_j} \right)^2 \right]^{1/2}}{(I)^{1/2}} \quad (3-46)$$

where the property $||W|| = \lambda_{\max}$ has been used. Equation (3-46) assumes that the noise is uncorrelated and with variance σ^2 . The value of R generated by Eq. (3-46) corresponds to the case where the vector \bar{f} is "directed" in the direction which produces the lowest possible $(S/N)_{\bar{f}}$.

If the noise is completely correlated, then R may be shown to be

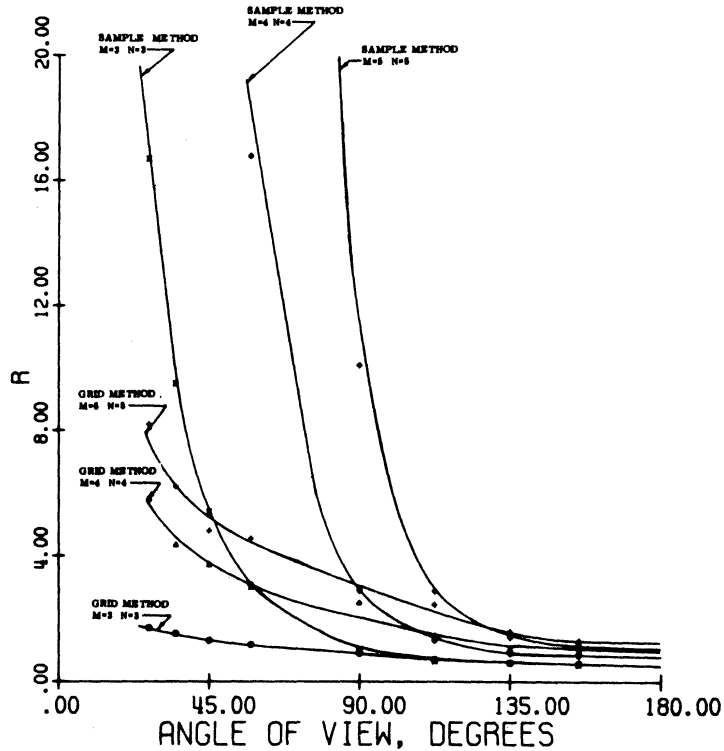
$$R \leq \lambda_{\max} / \lambda_{\min} \quad (3-47)$$

The value of R generated by Eq. (3-47) corresponds to the case where \bar{f} is "directed" in the direction which produces the lowest possible $(S/N)_{\bar{f}}$, while the noise is directed in the direction which produces the largest possible $(S/N)_{\phi}$. Equation (3-47) specifies the worst possible situation. The quantity defined by Eq. (3-47) is a function of the matrix \bar{W} only. In linear algebra, the ratio $\lambda_{\max} / \lambda_{\min}$ is called the condition of the matrix; a matrix with a large condition is termed nearly singular or poorly conditioned.

Figure 3-5a, b, and c are plots of the values of R generated by Eqs. (3-45), (3-46), and (3-47), as a function of the number of terms in the series expansions and the angle of view for both the sample and grid method. For Eq. (3-45), it is assumed that $f(x, y)$ is a radially symmetric Gaussian function (see Fig. 3-5b).

Figure 3-5 shows that the effects of noise are increased as either the resolution is increased or the angle of view is decreased. From Fig. 3-5, it is apparent that for a total angle of view of 90° or greater, the two techniques treat this idealized noise roughly the same. For a 45° angle of view, the grid method noticeably surpasses the sample method.

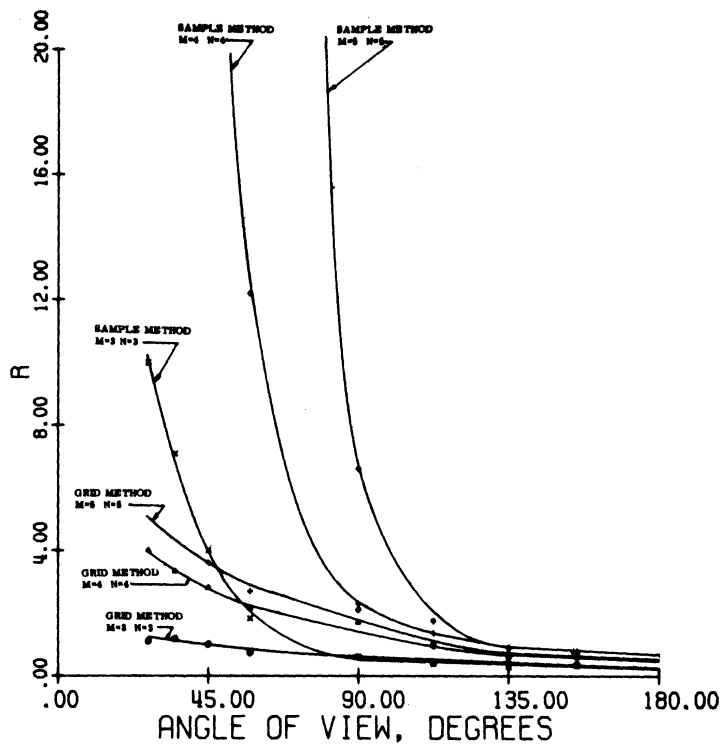
It must be recalled, however, that the sample method was introduced to minimize the input errors. The sample method generates a set of equations which is much more consistent than the grid method.



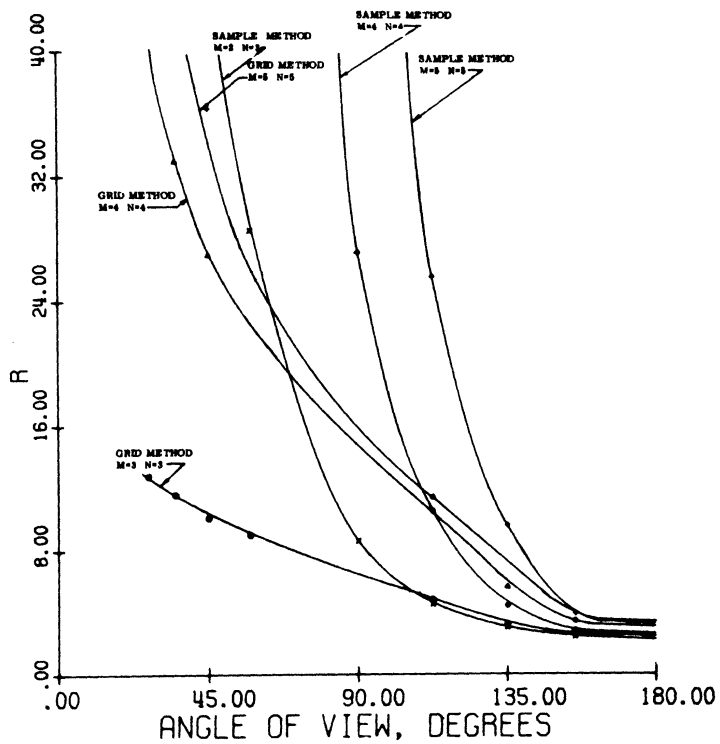
(a)

Fig. 3-5. Plot of Noise Amplification Factor vs. Angle of View and Resolution of Reconstruction for 10 Equally Spaced Views and 11 Rays/View

- (a) Assumes Uncorrelated and Identically Distributed Errors in the Pathlength Data, Generated by Eq. (3-45)
- (b) Assumes Uncorrelated and Identically Distributed Errors and Gaussian Input Distribution with $\sigma_x = \sigma_y = L_x/8 = L_y/8$, Generated by Eq. (3-46)
- (c) The Condition of the Matrix \bar{W} , Generated by Eq. (3-47)



(b)



(c)

Fig. 3-5. Plot of Noise Amplification Factor vs. Angle of View and Resolution of Reconstruction for 10 Equally Spaced Views and 11 Rays/View (continued)

It has been found from numerical simulations that, for index functions $f(x, y)$ which do not have severe discontinuities at the boundaries, the input errors of the sample method are three-to-ten times less than for the grid method. Therefore, the signal-to-noise ratio $(S/N)_{\bar{f}}$ for the sample method is three-to-ten times greater than for the grid method. Thus, the sample method normally generates a better reconstruction than the grid method unless the noise amplification factor R is three-to-ten times greater than the R factor generated by the grid method.

Since the success of the sample or the grid method depends on the nature of the input function, Table 3-1 was prepared to show a comparison between the sample and grid method for various input distributions $f(x, y)$. The table shows the results of reconstructions for three truncated Gaussian functions. Isometric views of the three Gaussian functions are shown at the top of Table 3-1. The parameters of the Gaussians have been adjusted so that the functions have differing discontinuities at the boundaries. The first Gaussian function, which is smoothly varying, is particularly sympathetic to the sample method while the third Gaussian, with large discontinuities at the boundaries, is particularly sympathetic to the grid method. The functions are reconstructed using 10 equally spaced views and 11 rays per view. These are the same parameters used to generate the plots shown in Fig. 3-5.

The table shows the following:

- (1) The norm of the residual vector defined as $||W_{ij}f_j - \phi_i||$.

Table 3-1 Comparison of the Sample Method and the Grid Method for Three Truncated Gaussian Functions With 10 Views and 11 Rays/View, M=4, and N=4.

Angle of View (degs.)	1*		2†		3‡	
	Sample Method	Grid Method	Sample Method	Grid Method	Sample Method	Grid Method
180	$ W_{ij}f_j - \phi_i $	18.1	4.0	14.2	9.6	4.8
	Error	0.9	1.1	2.5	4.6	1.0
	Condition	3.8	3.8	3.6	3.8	3.6
	R (Eqn. 3-46)	0.9	0.9	0.9	0.9	0.9
90	$ W_{ij}f_j - \phi_i $	13.4	2.6	6.7	5.5	2.6
	Error	1.4	1.1	3.7	3.6	4.8
	Condition	27.0	27.0	22.4	27.0	22.4
	R (Eqn. 3-46)	2.9	2.9	2.5	2.9	2.5
45	$ W_{ij}f_j - \phi_i $	16.7	4.2	9.0	9.0	3.2
	Error	12.1	3.5	8.0	26.8	6.9
	Condition	132.0	132.0	27.0	132.0	27.0
	R (Eqn. 3-46)	40.6	40.6	3.7	40.6	3.7

* $\sigma_x^2 = \sigma_y^2 = L_x/6 = L_y/6$

† $\sigma_x^2 = \sigma_y^2 = L_x/2 = L_y/2$

‡ $\sigma_x^2 = \sigma_y^2 = L_x/4 = L_y/4$

(2) The average error defined as

$$\text{error} = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N |f(m\lambda_x, n\lambda_y) - \bar{f}(m\lambda_x, n\lambda_y)|.$$

(3) The condition of the matrix W_{ij} defined by Eq. (3-47).

(4) The noise amplification factor R defined by Eq. (3-46).

The entries in the table indicate that, for the first Gaussian, the residual vector of the sample method is about three times smaller than the residual for the grid method. For the third Gaussian, with large discontinuities at the boundaries, the residual for the grid method is half the residual for the sample method.

The average errors listed in Table 3-1 show that the sample method provides more accurate reconstruction for five of the nine cases listed. For a 45° angle of view, the sample method nearly equals or surpasses the grid method for the first and second Gaussian functions. This is rather surprising since the noise factor for the sample method is 11 times greater than the noise factor for the grid method. It is also interesting to note that for a 90° angle of view the grid method provides the most accurate results for the first Gaussian and the least accurate for the third Gaussian. This is exactly the opposite of the results expected. This indicates that the effects of errors are a complicated process and are not completely described by the noise parameters shown in the table. The parameters do provide, however, an indication of the effects of errors on the reconstruction.

It should be noted that there is nothing that requires that the noise amplification factor R be less than one. Although an R factor of less than one implies that the noise level in the reconstruction is less than the noise level in the input signal, good reconstructions can still be obtained with an R greater than one, as shown in the table. In fact, one form of the noise factor, the condition, cannot be less than one (see Eq. (3-47)).

It is interesting to note that the noise amplification factor can be significantly decreased by reducing the number of resolution elements along the ray in the mean direction of view. For example, Fig. 3-5b shows a value of $R = 2.9$ for the sample method with a 90° angle of view, four resolution cells in the x direction and four resolution cells in the y direction, i.e., $M = 4$, $N = 4$. If the number of resolution cells along the mean direction of view is decreased to three, i.e., $M = 3$, then R will be reduced to 1.0. Reducing the number of resolution cells in this manner has the effect of increasing the effective angle of view, as defined by Eq. (3-11).

Using any of the series solutions developed in this chapter, it is necessary for the experimenter to select values of M and N which provide a reconstruction capable of resolving the relevant variations in the temperature field. In many temperature measurement experiments, the temperature field will vary much more rapidly in one direction than in the other. If this is the case, and a 180° angle of view is not available, the views that are available should be oriented in such a way that the variations along the ray in the mean direction

of view are minimized. This will maximize the effective angle of view and provide the lowest possible noise amplification factor.

Other Reconstruction Techniques

A procedure which can be used with any series expansion, referred to as the maximum entropy restoration, determines the coefficients of the series expansion so that the coefficients form a set with maximum entropy subject to the constraints imposed by the known path-lengths. This procedure represents an alternative approach to the least-mean-square criterion used to determine the series coefficients in the previous techniques.

Since entropy is not defined for negative numbers, it is necessary to assume that the refractive index field can be expressed as a point-wise positive function relative to some reference. This is a reasonable situation in many temperature measurement applications.

If the refractive index field is expanded in either a sample series, Eq. (3-22), or a grid series, Eq. (3-9), the result is a set of linear equations of the form

$$W_{ij} f_j = \phi_i \quad (3-48)$$

where W_{ij} is the weighing matrix with I rows and J columns; f_j is the vector whose elements are the coefficients in the series expansion.

Suppose that $I < J$; that is, the equation set is underdetermined.

There are an infinite number of solutions of Eq. (3-48) since there are $(J - I)$ independent variables. There is one solution of Eq. (3-48) which maximizes the entropy of f_j . This set will satisfy Eq. (3-48)

and be as smooth as possible. The entropy of f_j is

$$\text{entropy of } f_j = - \sum_{j=1}^J f_j \ln(f_j) \quad (3-49)$$

The objective is to maximize the entropy subject to the constraints imposed by the linear algebraic equation set of Eq. (3-48). This can be expressed as

$$(\text{entropy of } f_j) + \sum_{i=1}^I \lambda_i (\text{constraint equations}) = \max \quad (3-50)$$

where λ_i are undetermined Lagrange multipliers. Equation (3-50) can be expanded as

$$- \sum_{j=1}^J f_j \ln(f_j) + \sum_{i=1}^I \lambda_i (W_{ij} f_j - \phi_i) = \max \quad (3-51)$$

When the derivative of Eq. (3-51) is taken with respect to f_j for all values of j , and each term is set equal to zero, the result is

$$-1 - \ln(f_j) + \sum_{i=1}^I \lambda_i W_{ij} = 0 \quad (3-52)$$

or, solving for f_j explicitly,

$$f_j = \exp \left(-1 - \sum_{i=1}^I \lambda_i W_{ij} \right) \quad (3-53)$$

Equation (3-53) represents the basic result for f_j . Each value of f_j is an exponential function of the elements of the matrix W_{ij} .

It is necessary to determine the Lagrange multipliers which satisfy Eq. (3-48). Since the relationship between the Lagrange multipliers and the pathlengths is nonlinear, the Lagrange multipliers must be determined by an iterative procedure. This will constitute a formidable computation.

Each new ray introduces a new Lagrange multiplier and, hence, a new degree of freedom into Eq. (3-53). Thus, even though the refractive index field may be expanded in a series with many more terms than rays available, the final result will contain only as many degrees of freedom as independent rays used.

The effects of errors in the series expansion may make the technique unacceptable although the effects of noise on a superresolution scheme developed by Friedem [46], which is very similar to this technique, has been shown to be tolerable.

This technique has not been simulated on the computer because of the difficulties involved in solving the large number of simultaneous nonlinear equations that are generated.

Another technique which has been developed by researchers at State University of New York at Buffalo for applications in electron microscopy [31-33] is called Algebraic Reconstruction Techniques (ART). In interferometric applications, ART constitutes an iterative procedure which can be used to determine the coefficients of the series expansion of the refractive index field subject to the constraints imposed by the system of linear equations generated by the integration of the series expansion along the ray paths. The scheme

presented here is a modified form of the iterative method called the direct additive technique [31]. The original ART procedures have been developed to reconstruct internal structures from electron micrographs. The nature and amount of data available in electron microscopy makes it possible to develop the techniques in directions which provide greater resolution than is possible in interferometric applications in which only limited data is available. Thus, the results obtained by the simulations in electron microscopy and the results obtained here are not directly comparable.

The technique, as presented in Ref. 31, assumes that the function $f(x, y)$ is expanded in a series with many more terms than equations available. Thus, the equation set is underdetermined and there are an infinite number of solutions to the equation set. It has been proposed by Herman and co-workers (unpublished results) that this particular iterative procedure converges to the solution with minimum variance; that is,

$$||f_j|| = \min \quad (3-54)$$

The modified ART technique is developed here for the grid series expansion, Eq. (3-9). Let the linear equation set be

$$W_{ij} f_j = \phi_i \quad (3-55)$$

The technique requires that an initial guess be made for the vector f_j ; the elements of f_j are then corrected in an iterative procedure. If the value of f_j on the k th iteration is represented as $f_j^{(k)}$, then a measure of the inconsistency of the i th equation is

$$\Delta\phi_i = \phi_i - W_{ij}f_j^{(k)} \quad (3-56)$$

The pathlength error defined by Eq. (3-56) can be used to correct the elements of $f_j^{(k)}$ according to the formula

$$f_j^{(k+1)} = f_j^{(k)} + \frac{W_{ij}\Delta\phi_i}{\alpha_i} \quad (3-57)$$

where

$$\alpha_i = \sum_{j=1}^J (W_{ij})^2$$

The elements of f_j are corrected so that the i th equation is identically satisfied. The change in the j th element of the vector f_j is linearly proportional to the coefficient W_{ij} . This procedure is repeated for all of the ray equations. Unfortunately, as the values of f_j are altered to satisfy a particular equation, equations which have been previously satisfied will no longer be satisfied. Thus, the process must be repeated for the entire equation set; hopefully, after a number of iterations through the system of equations, the f_j vector will converge to some value which satisfies all of the equations.

The procedure was simulated numerically with interesting results. The number of degrees of freedom in the ART restoration is equal to the rank of the matrix W_{ij} . If all of the rays are independent, then each ray will introduce a new degree of freedom. Previous discussions have shown that redundant data is necessary to obtain good reconstructions. Numerical simulations confirmed that ART could not

reconstruct index fields without redundant data. Thus, it is necessary to obtain a matrix whose rank is less than the number of rays; i.e., some of the ray equations can be expressed as linear combinations of other rays. This can be accomplished by making the grid spacing larger than the average ray spacing. Another method would be to collect more pathlength measurements than terms in the series expansion. Collecting more rays than terms in the series expansion may produce an overdetermined system of equations. Although this iterative technique is apparently intended to be applied with underdetermined systems, numerical simulations showed that the primary requirement for acceptably accurate reconstructions is redundant data; overdetermined systems seemed to cause no problem.

During the iterative process, a parameter which is available to determine the convergence of the process is the norm of the residual. After the k th iteration, the norm of the residual is

$$R^{(k)} = \left\| \sum_j W_{ij} f_j^{(k)} - \phi_i \right\| \quad (3-58)$$

The residual cannot be zero unless the equation set is consistent. With redundant data, this is only possible if the series expansion converges point-wise to the actual refractive index field $f(x, y)$; this is rather unlikely in real applications. Thus, the iterative process can be terminated when

$$R^{(k+1)} - R^{(k)} \leq \epsilon \quad (3-59)$$

where ϵ is a small number. Several more iterations may also be executed to assure that the above condition remains satisfied.

It was found that the results of the reconstruction were a strong function of the last several ray equations satisfied using Eq. (3-57). To avoid this problem, a relaxation parameter was introduced into Eq. (3-57) so that

$$f_j^{(k+1)} = f_j^{(k)} + c \frac{W_{ij} \Delta \phi_i}{\alpha_i} \quad (3-60)$$

where c is a constant less than or equal to one. If c equals about 0.5 or less, at least during the last several iterative cycles, the final solution could be improved.

Like many iterative processes, it is difficult to predict the convergence of ART. Most of the analysis of iterative techniques that appear in the literature [47] define sufficient conditions for convergence. These conditions are not necessary for convergence, however, and tend to be very conservative.

The modified form of ART presented here can be used with any series expansion. From a computational viewpoint, the grid expansion is a particularly good choice since it generates a sparse matrix, although there is no systematic sparsity. If the sample method is used, it is probably better to solve the equation set using more direct procedures such as singular value decompositions. Unlike Golub's technique [42] which assumes that the matrix W_{ij} is full rank, a situation that does not necessarily exist here, singular value decomposition can be used to determine the least-mean-square solution to an underdetermined (or overdetermined) set of equations. The equation set is solved for the solution for which

$$\|f_j\| = \min$$

and

$$\|W_{ij}f_j - \phi_i\| = \min$$

The formal procedures will be superior to the iterative procedures for any series expansion so long as the computer time required to obtain the singular values is not too great. For applications in interferometry, the limited amount of data available generally restricts the number of unknowns to less than 100 so that the computer time is reasonable. (An Algol Program for singular value decomposition is given in Ref. 43, page 134.)

CHAPTER 4

COMPARISON OF THE RECONSTRUCTION TECHNIQUES

In the previous two chapters, a number of procedures to reconstruct a refractive index field from interferometric data have been developed or described. Six of these techniques have been programmed on the computer so that their performance could be compared for simulated temperature fields or used to reconstruct actual temperature fields from experimental data. The results of the comparison are presented here. The six techniques compared are:

- (1) Fourier Synthesis (Eq. 2-10)
- (2) Berry-Gibbs Integral (Eq. 2-16)
- (3) Grid Method (Eq. 3-9)
- (4) Sample Method (Eq. 3-22)
- (5) Frequency Plane Restoration (Eq. 3-26)
- (6) Algebraic Reconstruction Technique (ART) (Eq. 3-57)

The simulation uses the following procedure. A hypothetical temperature field is specified for a particular fluid. The refractive index field corresponding to this temperature field is numerically integrated along typical ray paths. The integral values, representing the pathlength measurements, provide the input data necessary for the various reconstruction procedures. The procedures then reconstruct the refractive index field without reference to the original function.

A discrete amount of pathlength data is used in the simulations. The amount of data used roughly corresponds to the amount of data that would be available in a typical experimental application.

To include the effects of refraction, a ray-tracing routine was written to define the paths of integration. For the temperature fields simulated here, refraction effects were found to be minor; to avoid the time consuming ray-tracing operation, the refractive index field was merely integrated along straight lines.

The reconstructed temperature fields are compared on a point-wise basis at 81 uniformly spaced locations within the temperature field.* The 81 locations consist of nine rows of points in the y direction, distributed uniformly between $y/L_y = 0.1$ and $y/L_y = 0.9$, and nine columns of points in the x direction between $x/L_x = 0.1$ and $x/L_x = 0.9$. The error at a point i is defined as

$$e_i = \left(\frac{T_{\text{act}} - T_{\text{rec}}}{\Delta T_{\text{max}}} \right)_i$$

where T is the temperature and T_{act} and T_{rec} represent the actual and reconstructed temperatures, respectively. The average error is defined as

$$e_{\text{av}} = \frac{1}{81} \sum_{i=1}^{81} |e_i|$$

The results of the simulations are presented here for three different asymmetrical temperature fields. Figures 4-1a, 4-2a, and

*To allow for comparisons at arbitrary points, an interpolation formula of the same form as Eq. (3-13) is used to interpolate between elements in the grid and ART methods.

4-3a show isometric plots of the simulated temperature fields which were used. The three input functions selected are continuous and smooth, except for small discontinuities at the boundaries in some cases. Functions of this form are used because they are similar to the types of distributions that will be found in real interferometric studies of temperature or other diffusion-type fields which yield smooth variations. As discussed in the previous chapters, the accuracy of the various techniques depends on the characteristics of the input function. In this sense, numerical simulations for a particular input function do not provide a completely general comparison of the techniques. These simulations do, however, represent a good cross section of the types of temperature fields likely to be found in actual applications.

Tables 4-1, 4-2, and 4-3 summarize the results of the comparisons. In all of the comparisons, pathlength data was collected at 19 equally spaced positions for each of the 15 equally spaced directions of view so that the various reconstructions could be compared. The tables also show the computer time (CPU seconds on an IBM-360(67)) necessary to reconstruct the temperature field from the pathlength data; input/output times and other auxiliary operations are not included. Where relevant, the number of terms used in a series expansion is listed in the table. Figures 4-1b, 4-2b, and 4-3b show isometric plots of typical reconstructions using the techniques listed in the figure captions.

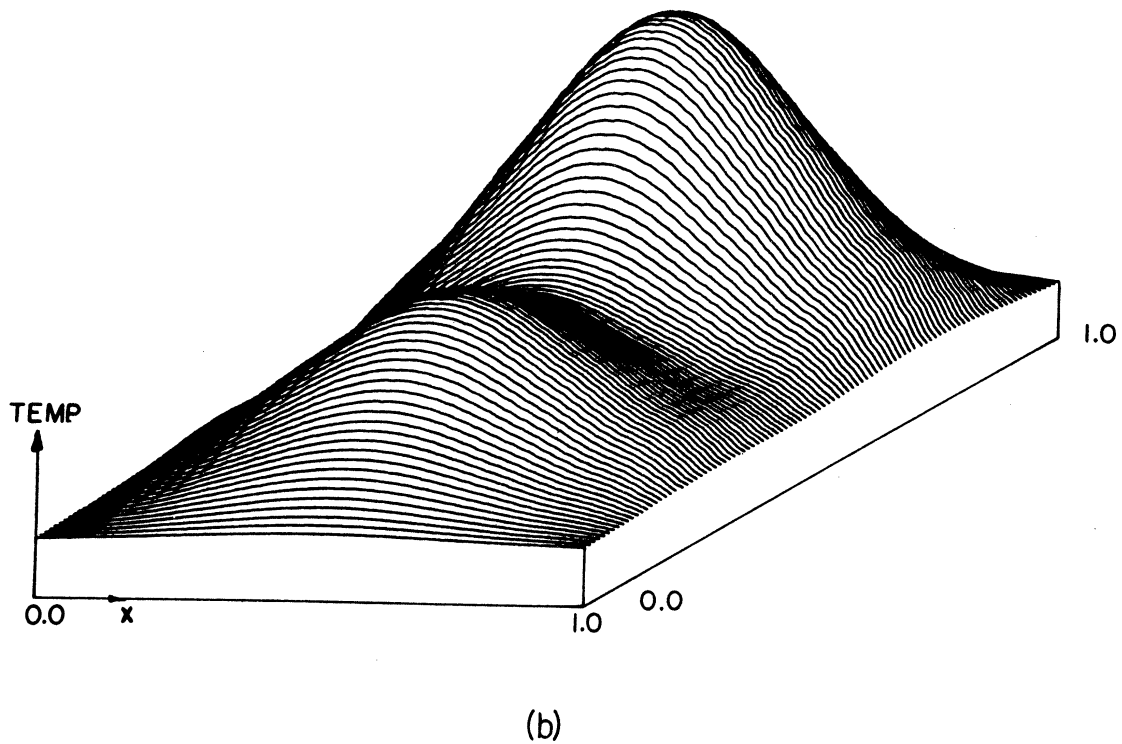
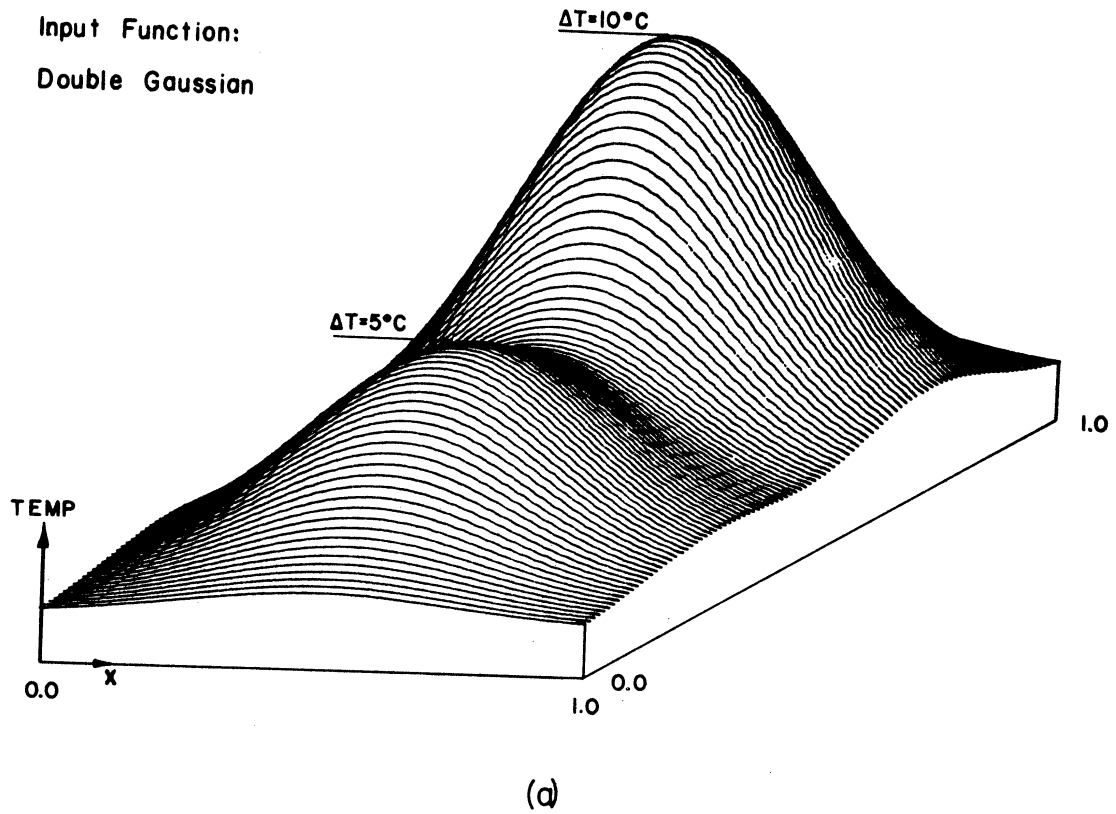


Fig. 4-1. Isometric Plot of Double-Gaussian Function

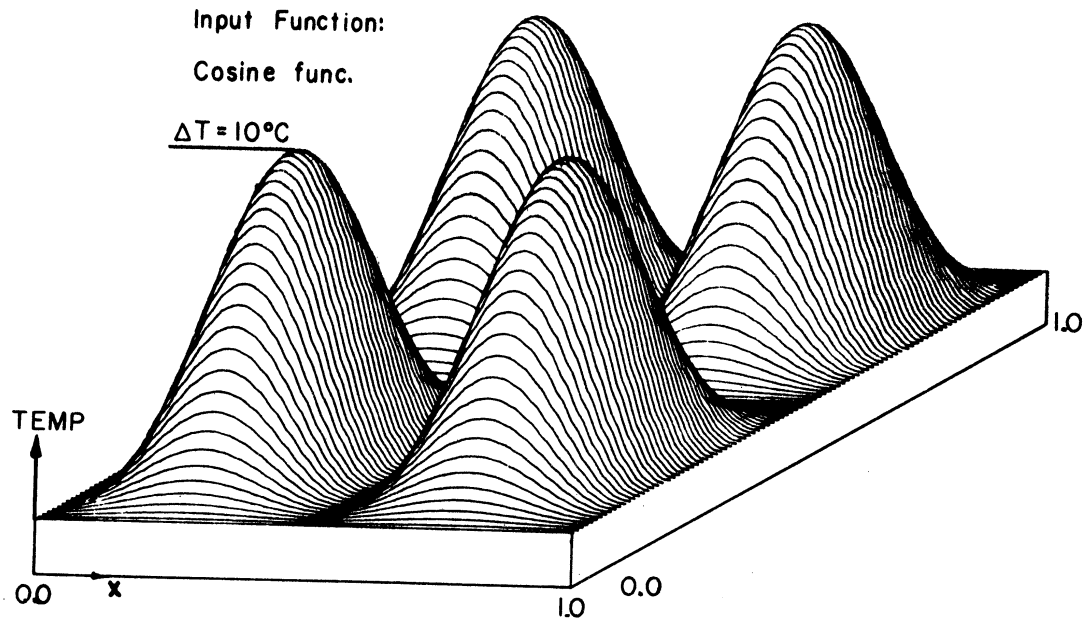
- (a) Input Function
- (b) Reconstruction by Frequency Plane Restoration Technique with 45° Angle of View (see Table 4-1 for other parameters)

Table 4-1 Comparison of Results

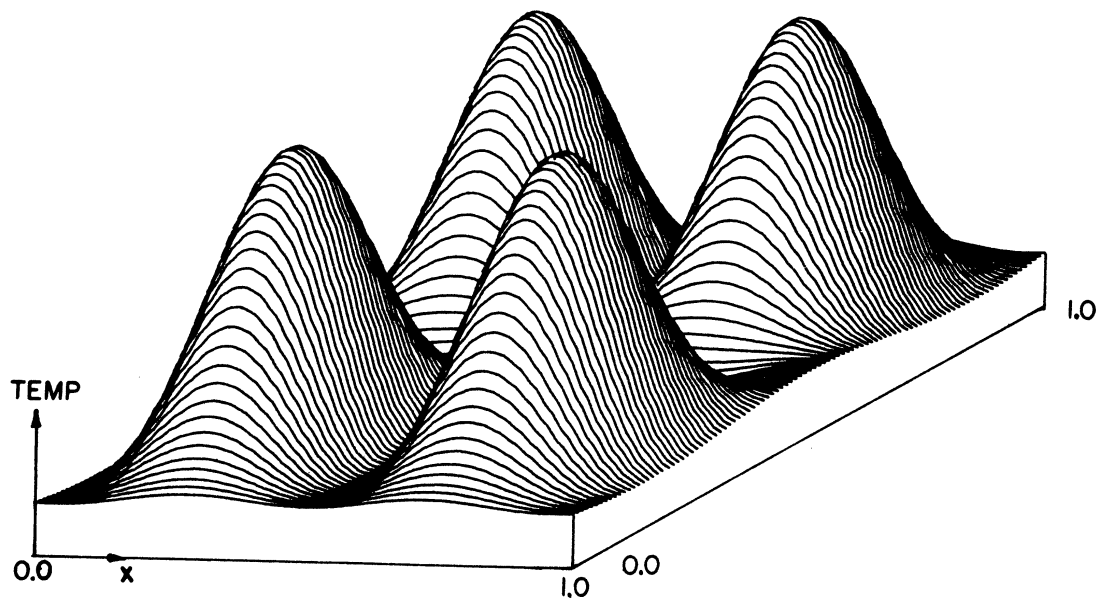
Input Function: Double Gaussian

$$T(x,y) = T_0 + \alpha \left\{ \Delta T_1 \exp[-8.(x-.5)^2 - 25.(y-.14)^2] + \Delta T_2 \exp[-8.(y-.5)^2 - 25.(y-.72)^2] \right\}$$

Method	Number of Views	Rays View	Angle of View (degrees)	Average Error (%)	Maximum Error (%)	Computer Time (IBM 360-67) CPU secs.	Comments
Fourier Synthesis	16	19	180	1.63	4.32	3.4	
Berry-Gibbs Integral	16	19	180	0.67	1.45	2.5	
Sample Method	16	19	180	0.30	1.06	16.7	M = 7, N = 7
				0.87	3.15	10.3	M = 5, N = 7
				5.44	12.33	2.8	M = 3, N = 5
Grid Method	16	19	180	1.43	5.21	11.0	M = 7, N = 7
				2.10	10.59	6.5	M = 5, N = 7
				2.61	8.73	1.4	M = 3, N = 5
Frequency Plane Restoration	16	19	180	0.66	1.82	16.1	M = 7, N = 7
				1.09	2.69	10.5	M = 5, N = 7
				0.70	1.86	3.2	M = 3, N = 5
Art Technique	16	19	180	1.24	3.29	15.4	M = 9, N = 9
				2.99	12.82	12.2	M = 6, N = 9
				4.70	14.92	7.0	M = 4, N = 6



(a)



(b)

Fig. 4-2. Isometric Plot of Cosine Function

(a) Input Function

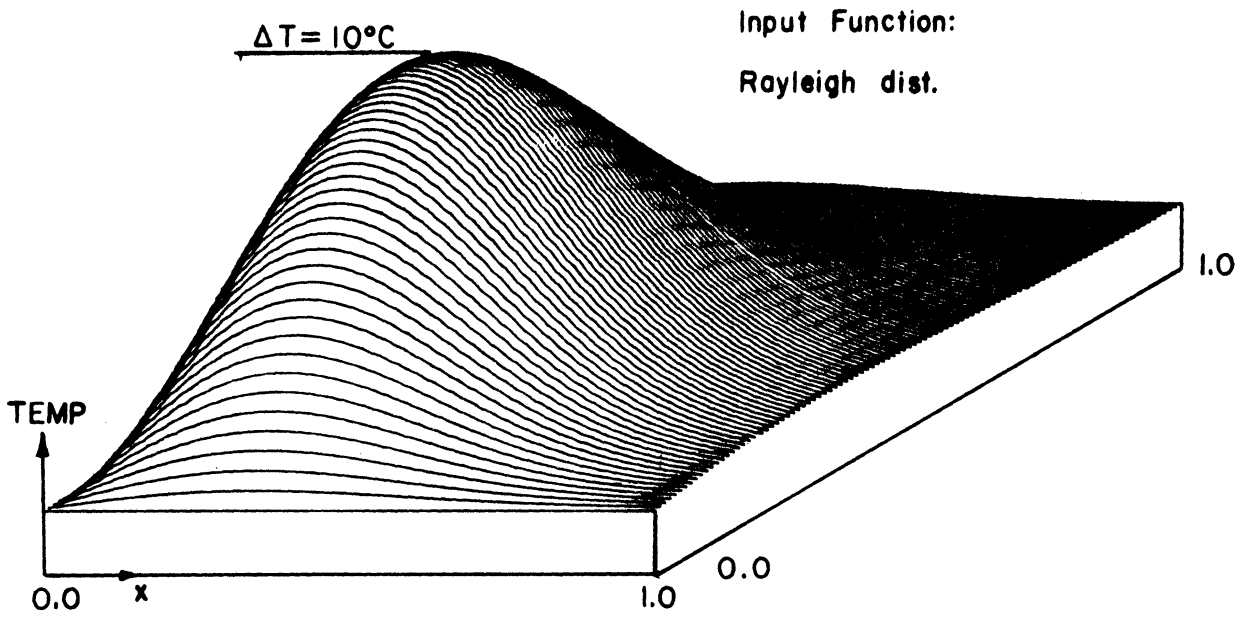
(b) Reconstruction by Sample Method with
 90° Angle of View (see Table 4-2
for other parameters)

Table 4-2 Comparison of Results

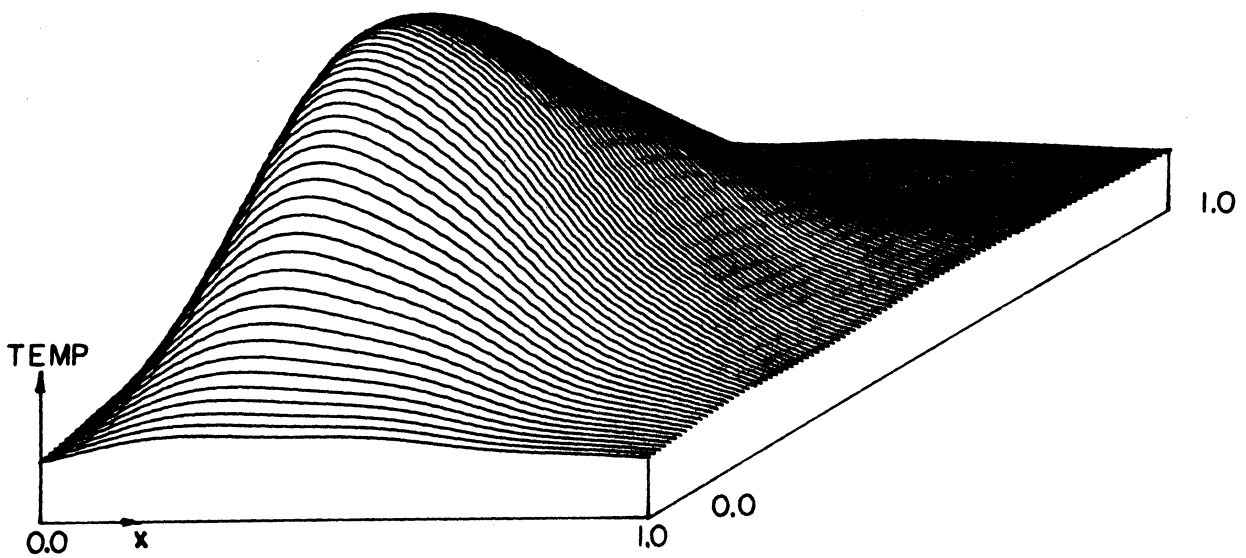
Input Function: Cosine func.

$$T(x,y) = T_0 + \alpha \Delta T / 4 [1 - \cos \{4\pi(x-.5)\}] [1 - \cos \{4\pi(y-.5)\}]$$

Method	Number of Views	Rays View	Angle of View (degrees)	Average Error (%)	Maximum Error (%)	Computer Time (IBM 360-67) CPU secs.	Comments
Fourier Synthesis	16	19	180	3.08	7.76	3.3	
Berry-Gibbs Integral	16	19	180	1.03	3.13	2.5	
Sample Method	16	19	180 90	0.28 2.23	0.79 8.18	16.5 10.3	M = 7, N = 7 M = 5, N = 7
Grid Method	16	19	180 90	3.10 22.70	12.71 72.3	11.3 6.5	M = 7, N = 7 M = 5, N = 7
Frequency Plane Restoration	16	19	180 90	0.22 0.24	1.06 1.17	17.4 10.2	M = 7, N = 7 M = 5, N = 7
Art Technique	16	19	180 90	2.21 4.76	8.21 15.22	15.8 11.3	M = 9, N = 9 M = 6, N = 8



(a)



(b)

Fig. 4-3. Isometric Plot of Rayleigh Distribution

(a) Input Function

(b) Reconstruction by Grid Method with
180° Angle of View (see Table 4-3
for other parameters)

Table 4-3 Comparison of Results

Input Function: Rayleigh dist.

$$T(x,y) = T_0 + \alpha \Delta T xy (9\sqrt{e}) \exp[-4.5(x^2 + y^2)]$$

Method	Number of Views	Rays View	Angle of View (degrees)	Average Error (%)	Maximum Error (%)	Computer Time(IBM 360-67) cpu secs.	Comments
Fourier Synthesis	16	19	180	1.52	4.65	3.4	
Berry-Gibbs Integral	16	19	180	0.84	2.18	2.5	
Sample Method	16	19	180	1.08	4.45	16.8	M = 7, N = 7
			90	2.40	9.12	9.7	M = 5, N = 7
			45	9.53	27.41	2.8	M = 3, N = 5
Grid Method	16	19	180	1.40	5.33	11.2	M = 7, N = 7
			90	3.95	13.34	6.2	M = 5, N = 7
			45	4.93	15.02	1.5	M = 3, N = 5
Frequency Plane Restoration	16	19	180	1.07	2.76	16.6	M = 7, N = 7
			90	1.74	5.87	10.3	M = 5, N = 7
			45	5.47	16.70	3.2	M = 3, N = 5
Art Technique	16	19	180	1.19	4.35	9.0	M = 9, N = 9
			90	3.77	12.93	9.7	M = 5, N = 7
			45	5.95	15.10	7.3	M = 4, N = 6

The fluid simulated in these numerical experiments is water. The relationship between temperature and refractive index [48] at an ambient temperature of 23°C and a wavelength of 6328Å is approximately linear in the temperature ranges involved in these simulations and is given by

$$(n - n_o) = +\alpha(T - T_{amb})$$

where $\alpha = -1.138 \times 10^{-4}$ and T is the temperature in degrees centigrade.

Based on the results of these comparisons, several conclusions can be drawn. If a 180° angle of view is available, all of the techniques produce good reconstructions with a typical average error of about 1%. Three of the techniques are particularly accurate: the sample method, the Berry-Gibbs integral, and the frequency plane restoration method. Since the Berry-Gibbs integral does not require the experimenter to select the number of terms in a series expansion and since the procedure requires considerably less computer time, the Berry-Gibbs integral technique is generally the best choice.

If the angle of view is less than 180°, two of the techniques again produce the most reliable results: the sample method and the frequency plane restoration technique. For a 45° angle of view, the frequency plane restoration procedure produces the most accurate reconstructions.

To obtain the results in the numerical simulations of the three fields just presented, pathlength data is collected for views which are centered about the x axis. For the double-Gaussian function shown in Fig. 4-1, this is the optimum direction of observation.

For the other two simulated functions, the mean direction of view is of little consequence. Table 4-4 compares the effect of collecting data for views centered about the y axis or the x axis for the double-Gaussian function. For a 90° angle of view, the error is not increased appreciably. This is because the five resolution elements along the mean direction of view are sufficient to accurately represent the input function. For a 45° angle of view, the errors are increased significantly if the mean direction of view is along the y axis. The three resolution elements along the mean direction of view are not sufficient to resolve the double-Gaussian function.

Table 4-4 Comparison of the Sample Method and the Frequency Plane Restoration Technique as a Function of the Mean Direction of View for the Double Gaussian Function Shown in Figure 4-1.

Angle of View (degrees)	FREQUENCY PLANE RESTORATION		SAMPLE METHOD		
	Average Error (per cent)	Maximum Error (per cent)	Average Error (per cent)	Maximum Error (per cent)	
180	0.6	1.8	0.3	1.1	M = 7 N = 7
90 about x axis	1.1	2.7	0.9	3.2	M = 5 N = 7
90 about y axis	0.8	2.5	0.9	3.5	M = 7 N = 5
45 about x axis	0.7	1.9	5.4	12.3	M = 3 N = 5
45 about y axis	9.4	29.7	8.2	25.2	M = 5 N = 3

CHAPTER 5
EXPERIMENTAL APPLICATIONS TO
TEMPERATURE MEASUREMENT

A genuine indication of the applicability of the reconstruction schemes developed thus far can only be obtained by conducting an actual experiment. In this chapter, the results of two experiments are presented in which the temperature fields in the steady-state convective plumes are reconstructed using holographic pathlength data.

Experimental Considerations

Before presenting the results of the experiments, several practical considerations dealing with the reduction and interpretation of the pathlength data from an interferogram will be discussed. One effect encountered in holographic interferometry that can seriously interfere with the reduction of the interferometric data is coherent speckle. This effect is an inevitable result of using highly coherent diffuse illumination. Basically, coherent speckle is an interference phenomena which results in a random "salt and pepper" noise pattern being superimposed on the interferogram. This is very detrimental since the visibility of the interferometric fringes rapidly decrease as the fringe frequency becomes of the order of the characteristic speckle frequency.

The problem of fringe clarity has been considered in some detail by Tanner [49, 50]. Utilizing Goldfischer's [51] model for coherent speckle, he redefines fringe visibility in a manner which is appropriate for characterizing interferometric fringe in the presence of speckle. This definition is then applied to the particular case of holographic interferometry with a diffuse light source and the optimum viewing aperture size is determined. As an example, it is shown that for a typical object having an axial dimension of 10 cm, sufficient fringe clarity would be obtained only for optical path-length changes not exceeding about one fringe per millimeter. In most applications, this is unacceptable.

The problem with coherent speckle can be overcome if one is willing to view the interferogram from a number of discrete viewing angles rather than completely arbitrary directions. This is perfectly reasonable since all of the reduction techniques involve data collected from a finite number of viewing angles. A technique for achieving this is discussed in detail by Vest and Sweeney [52]. Basically, the diffusing screen used to backlight the transparent object is replaced by a phase grating. When a sinusoidal phase grating of frequency f is illuminated by a plane wave of unit amplitude, the complex amplitude of the light field immediately to the right of the grating is given by

$$u(x, y) = \sum_{n=-\infty}^{+\infty} J_n\left(\frac{m}{2}\right) \exp(i2\pi n f_x)$$

where m is the peak-to-peak phase delay and the finite aperture of the grating is neglected. Hence, the energy of the incident wave is deflected into a number of component plane waves propagating at angles of $\pm n\lambda$ with respect to the normal. The interference pattern for each discrete viewing angle can be obtained by optically Fourier transforming the holographic reconstruction and sampling only in the neighborhood of a single spectral component in the transform plane. The corresponding portion of the reconstruction can then be imaged or processed in any desired manner. It has all the properties associated with an interferogram produced with a single collimated object beam.

Experiments have shown that the fringe resolution can easily be increased five-fold over that obtained with diffuse illumination. To avoid undesirable overlap (i.e., aliasing) between the diffracted orders, the maximum fringe frequency cannot exceed half the fringe frequency of the grating.

Another requirement for accurate reconstructions of the temperature field is that the pathlength data be collected at a rate proportional to the effective bandwidth of the refractive index field. If the center of the light and dark lines on the interferogram are used as data points, then the pathlength is known for positions on the pathlength function $\phi(\rho, \theta)$ where

$$\phi(\rho, \theta) = \frac{m\lambda}{2}$$

where m is an integer and λ is the wavelength of the light. Figure 5-1 shows a typical pathlength function; the positions corresponding

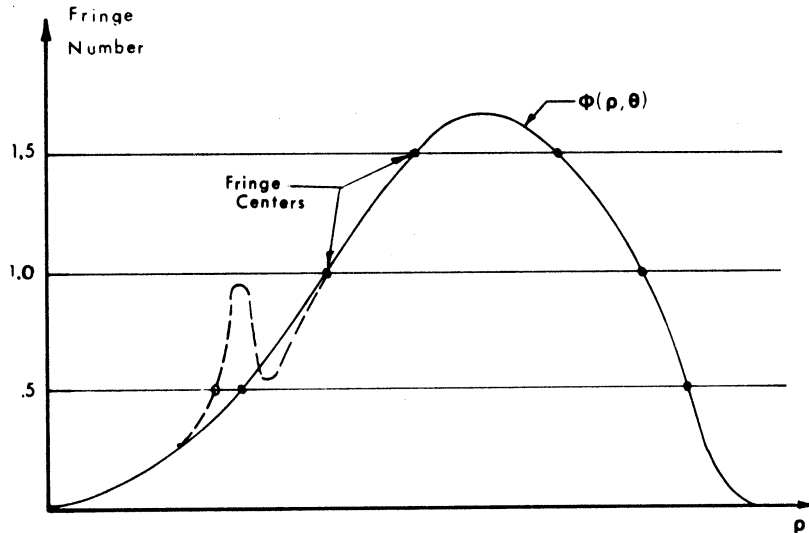


Fig. 5-1. Plot of Typical Pathlength Function Showing Positions of Center of Fringes and the Possible Effect of Aliasing

to the centers of the light and dark fringes are labeled. If $\phi(\rho, \theta)$ contains high spatial frequency terms whose amplitudes are not sufficient to produce interferometric fringes, the resultant samples will contain aliasing effects. Thus, the sampling rate required by the reconstruction techniques of twice the effective bandwidth of the refractive index field may not be compatible with the number of data points available. Theoretically, this can be overcome by using a photodetector to record the intensity of the interferogram so that fractional fringes may be resolved. This is often done in classical

two-dimensional interferometric systems; but in holographic interferometry using diffuse illumination, the noise introduced by coherent speckle is too high to warrant such a procedure. These sampling problems can be overcome by sufficiently increasing the amplitude of the refractive index field, by decreasing the wavelength of the light used, or by using phase grating illumination to eliminate coherent speckle and a photodetector to resolve fractional fringes. In some cases, none of these techniques will be satisfactory and the fringe resolution stands as an experimental limitation.

All of the techniques developed to reconstruct the three-dimensional temperature fields have neglected the effect of ray curvature due to refraction. Although refraction cannot be altered by changes in the experimental system, its effect can be minimized by use of a proper imaging system to view the fringes in an optimum manner.

Figure 5-2 shows the effect of neglecting refraction. A ray traversing the refractive index field enters at point A, crosses the refractive index field to point B, and continues along a straight line \overline{BC} to the imaging system. The hologram is generally inserted to the left of the imaging system and will introduce a slight distortion in the interferogram. This effect is very minor, however, and will be neglected here. When viewed through the imaging system, the apparent position of the ray is the intersection of the extension of the line \overline{BC} and the focal plane of the imaging system. If refraction is neglected, the pathlength of the ray \overline{ABC} is assigned to

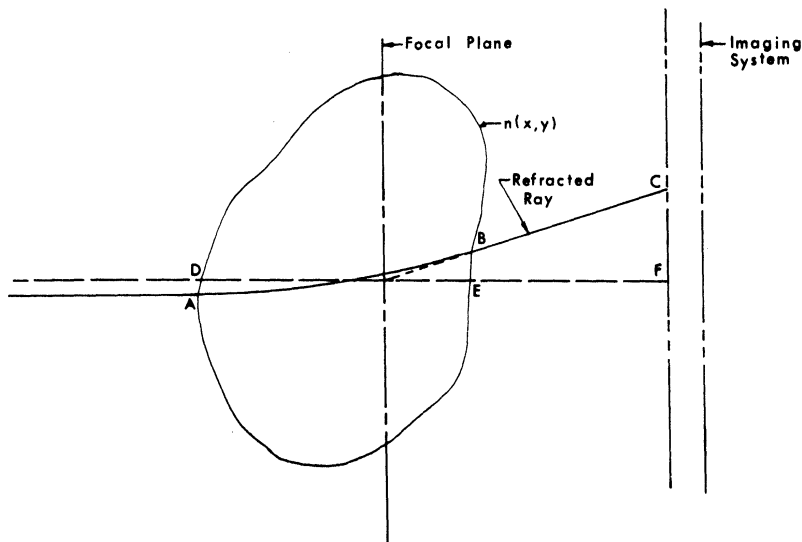


Fig. 5-2. The Effect of Neglecting Refraction,
the Pathlength of Ray ABC is Assigned to the
Hypothetical Straight Ray DEF

the hypothetical straight ray \overline{DEF} . In two-dimensional interferometry, a first-order correction for the effects of refraction can be made if the basic functional form of the refractive index distribution is known (cf. [1]). The corrections can be introduced by proper positioning of the focal plane of the viewing system. Such a procedure could be introduced into three-dimensional interferometry if the basic form of the refractive index field is known or possibly determined using an iterative process. With no a priori information available,

however, the focal plane should be located in approximately the center of the refractive index field since this would tend to minimize the fringe shifts introduced by ray curvature.

Temperature Above a Short Heated Wire

As an experimental demonstration of the sample method for reconstructing temperatures, the temperature distribution in a steady-state thermal plume above a short heated wire was measured by holographic interferometry. Figure 5-3 is a photograph of the wire and its support. The tungsten wire is 2.0 cm long and is supported

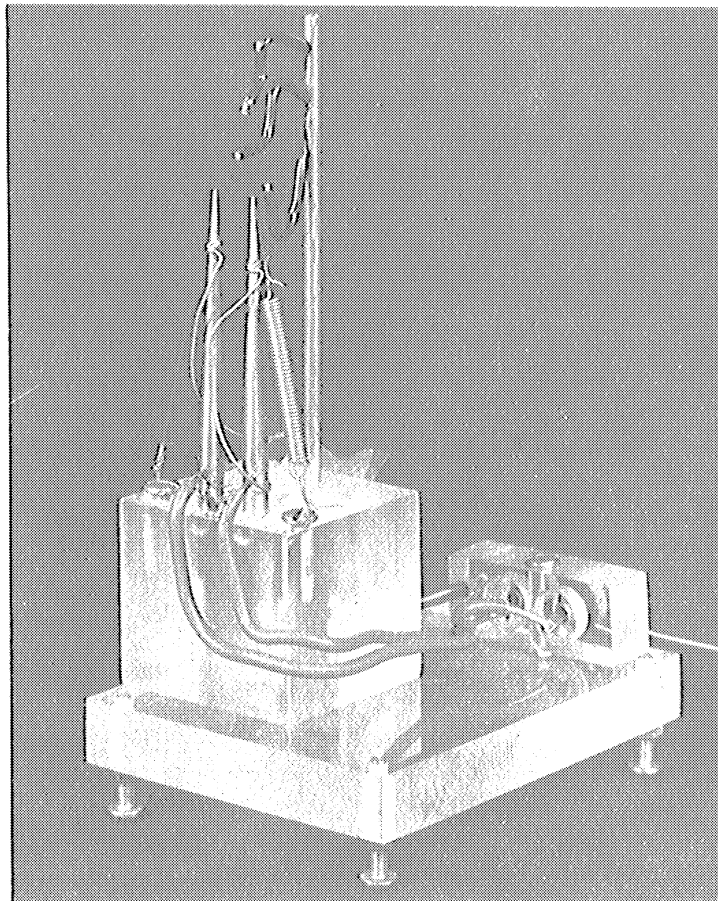


Fig. 5-3. Photograph of the Short Heated Wire and Its Support Stand

between two slender copper posts. Also visible in Fig. 5-3 are the thermocouple probes used to verify the interferometric temperature measurements.

The wire was submerged in water and heated electrically between the two support posts at a rate of 1.9 watts/cm. An interferogram was formed by making two holographic exposures, one while the wire was at ambient temperature and the other while it was being heated at a steady rate.

A schematic diagram of the experimental apparatus is shown in Fig. 5-4. The laser beam was divided into a spherical reference

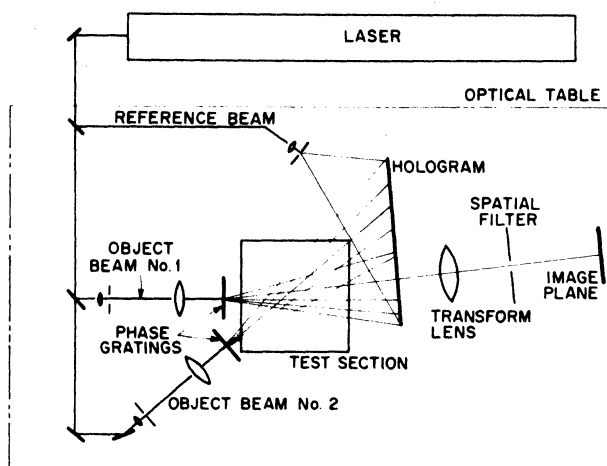


Fig. 5-4. Schematic of the Experimental System Used in the Heated Wire Experiments, the Object Illumination is Derived from Phase Gratings

wave and two object waves. Each object wave was filtered, expanded, collimated, and then passed through a phase grating oriented to its direction of propagation. The phase gratings had spatial frequencies of 100 lines/mm. The total wavefront emerging from the test section was recorded holographically on a 30.5-cm long Agfa 10E70 photographic plate. To avoid the inconvenience of arranging this apparatus within the constraint imposed by a short coherence length, a mode selector was added to the He-Ne laser used in this experiment. The device used for selection of a single mode was first proposed and demonstrated by Currie [53]. The device, which consists primarily of a low-reflectivity mirror inserted into the laser cavity, achieves single-mode operation with a power loss on the order of 10%.

The optical apparatus enabled interferograms to be recorded for eight different directions within a total angle of view of 20.5° . Since the temperature field is symmetric about the vertical plane containing the wire, this is equivalent to a 41° range of viewing angle. Figure 5-5 shows a typical interferogram produced by this experiment. Two thermocouples, which were used to verify the interferometric temperature measurement, can also be seen in Fig. 5-5.

It was a straightforward matter to obtain the data necessary to deduce the temperature in a given horizontal plane. The angle of view for each interferogram (e.g., that shown in Fig. 5-5) was determined by measuring the apparent spacing of the two vertical pins which support the horizontal wire; then the position and order of each fringe was noted. The values of $\phi(\rho, \theta)$ required for the

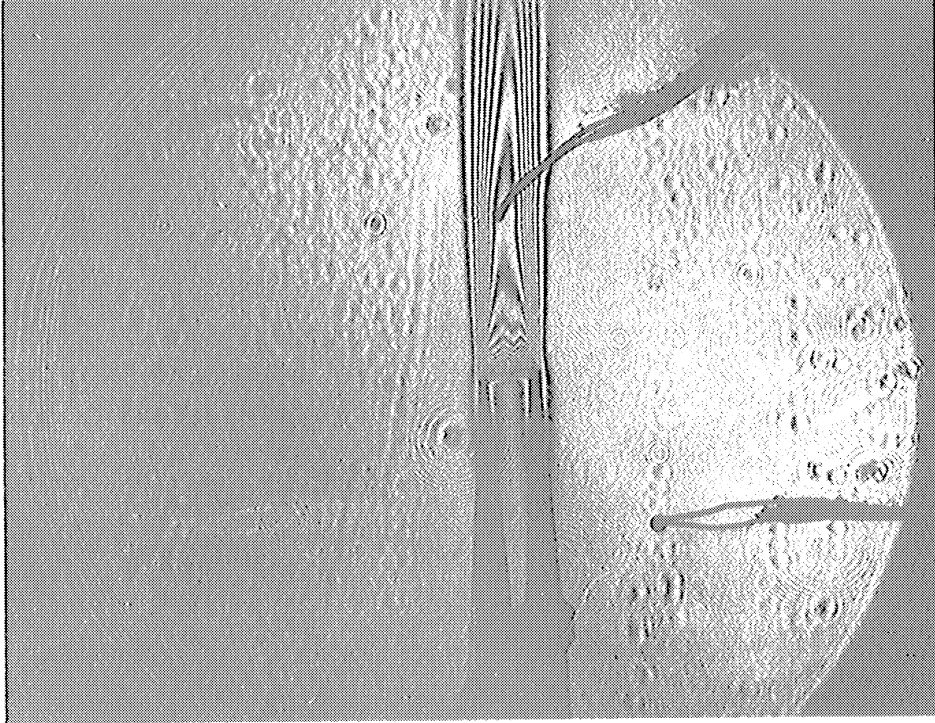


Fig. 5-5. Typical View of an Interferogram of the Short Heated Wire in Water

analysis are simply the product of these fringe-order numbers and the wavelength of the laser light. It should be noted that this process of data retrieval is greatly simplified by the use of an overdetermined system since it is not necessary to specify a priori the ray directions for which $\phi(\rho, \theta)$ must be determined.

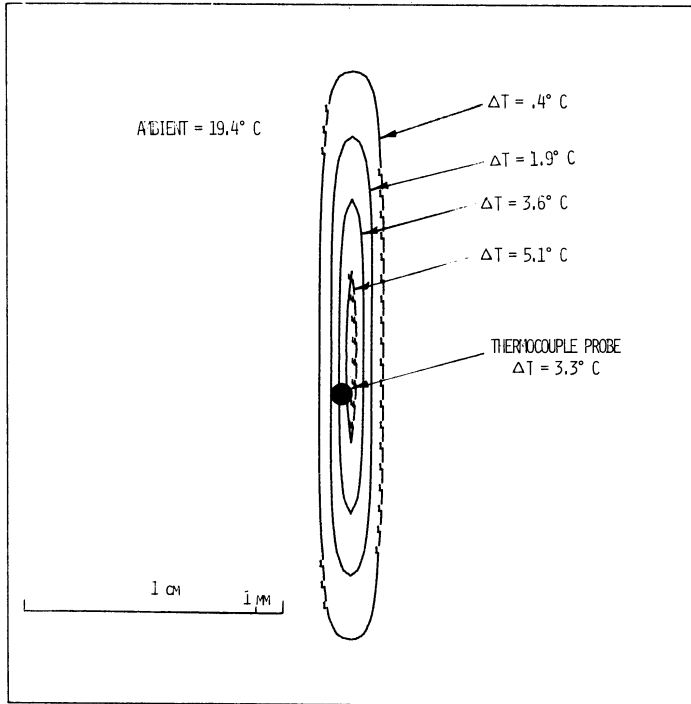
The data obtained in this manner was introduced into Eq. (3-22). The resulting system of equations was solved by the least-mean-squares

method to yield $\bar{f}(x, y) = [n(x, y) - n_0]$, where n_0 is the refractive index of the water far from the heated wire. Using the known functional relationship between refractive index and temperature [48], the temperature field, $T(x, y)$, was determined in a horizontal plane slightly below the upper thermocouple in Fig. 5-5. Isothermal contours obtained with this procedure are shown in Fig. 5-6a and an isometric plot of the temperature in this plane is shown in Fig. 5-6b. The thermocouple inserted into the plane was used to obtain an independent measurement of the temperature. It indicated a temperature rise of 3.3°C above the ambient temperature. Although the thermocouple junction was only 0.7 mm in diameter, it occupied the space between the isotherms of 2.0°C and 5.6°C , as shown in Fig. 5-6a. Because the thermocouple provides only an average temperature over this region, a precise measure of the error in the interferometric technique was not obtained; however, the results are clearly reasonable.

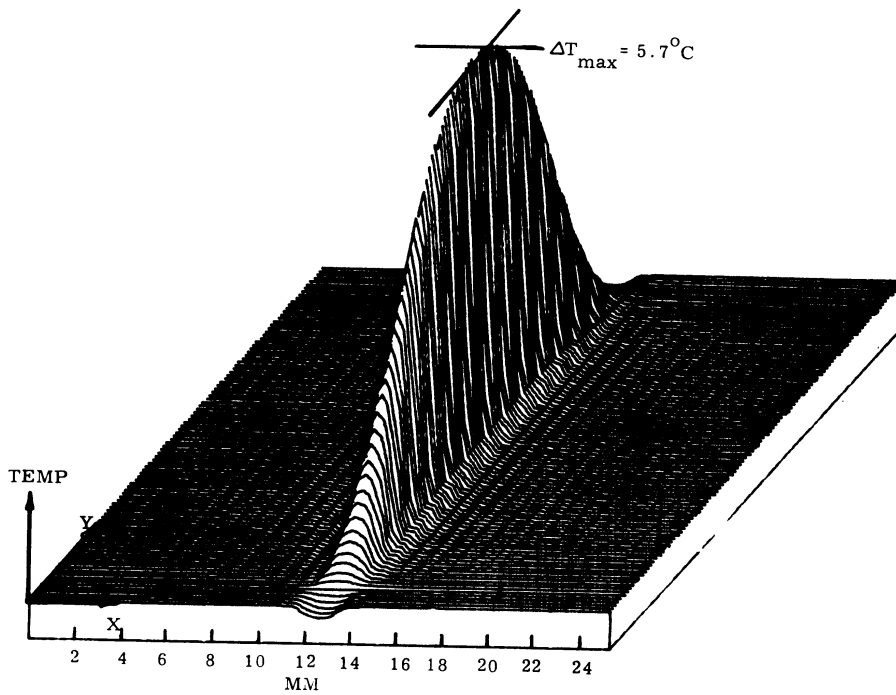
Temperature Above a Horizontal Heated Surface

The objective of these experiments was to reconstruct the temperature field in the steady-state thermal plume above a horizontal rectangular surface heated electrically in water.

Figure 5-7 is a photograph of the heating element and its support stand. The heated surface was fabricated from two aluminum plates approximately 2.5 cm long by 1.9 cm wide and 0.3 cm thick. A 2.3-cm long section of heating tape was sandwiched between the two aluminum plates. Black masking tape enclosed the edges of the two aluminum plates; the tape extended approximately 1 cm below the bottom of the



(a)



(b)

Fig. 5-6. Reconstructed Temperature Field in a Plane Above the Heated Wire

- (a) Isothermal Contour Plot of Reconstructed Temperature
- (b) Isometric Plot of Reconstructed Temperature

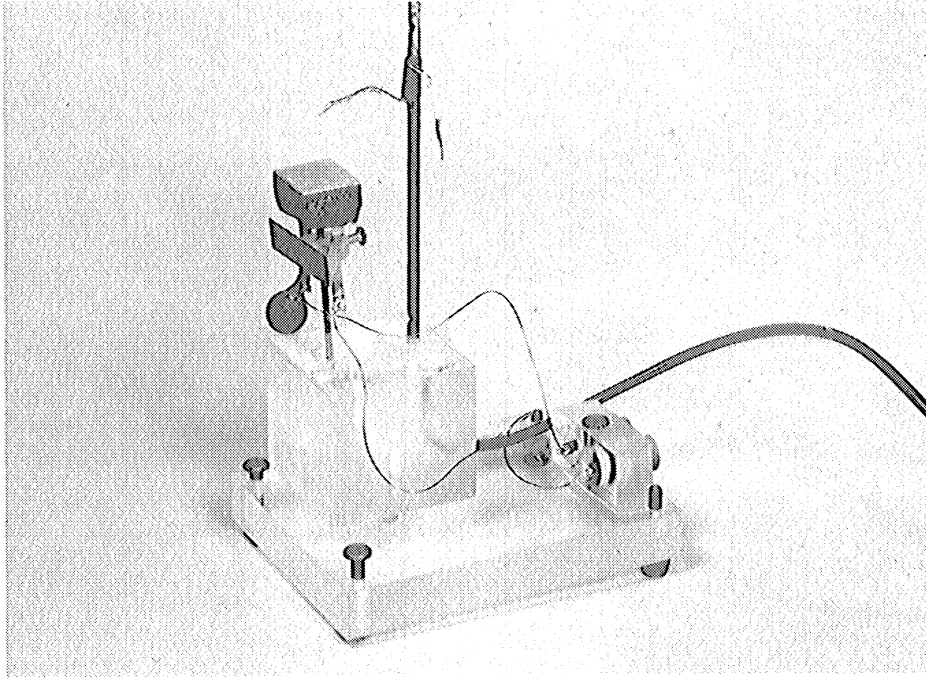


Fig. 5-7. Photograph of the Heated Plate and Its Support Stand

two aluminum plates. Also visible in the photograph is the thermocouple probe used to verify the reconstructed temperature. Below the aluminum plates is a 1.2-cm diameter sphere used as a dimensional reference. Slightly above the sphere is a vertical plate which also served as a dimensional reference for determining the viewing direction. By measuring the distortions in the image of the sphere, it was possible to compensate for the refractive distortions caused by the components of the test chamber.

The entire test fixture shown in Fig. 5-7 was submerged in a glass-walled tank 30 cm long and 15 cm wide, filled to a height of approximately 20 cm with distilled water. The heated surface was approximately 11 cm from the bottom of the tank.

The optical system used in these experiments was essentially the same as that shown in Fig. 5-4. In this experiment, however, the fringe frequency was low enough to allow diffuse illumination to be used. The phase gratings shown in Fig. 5-4 were replaced with a high-quality opal glass diffuser. The diffuser was illuminated with a single 13-cm diameter collimated beam. Data could be collected over an angle of about 30° which, due to symmetry about the center of the plate, was equivalent to a total viewing angle of 60° . The interferometric data was recorded directly from the interferogram by observing the interferogram through an imaging lens and a calibrated eyepiece mounted on a translation stand.

The hologram was recorded on a 30.5-cm long Agfa 10E70 photographic plate. An interferogram was formed by making two holographic exposures, one while the fluid was at the ambient temperature and the other exposure while the plate was being heated. Figure 5-8 shows a typical view of the interferogram. The plate was heated at a rate of 1 watt/cm^2 . The pathlength data was collected so that the temperature field could be reconstructed on two horizontal planes above the plate, which were 0.5 cm and 1.0 cm above the heated surface, respectively. The positions of the planes are shown in Fig. 5-8.

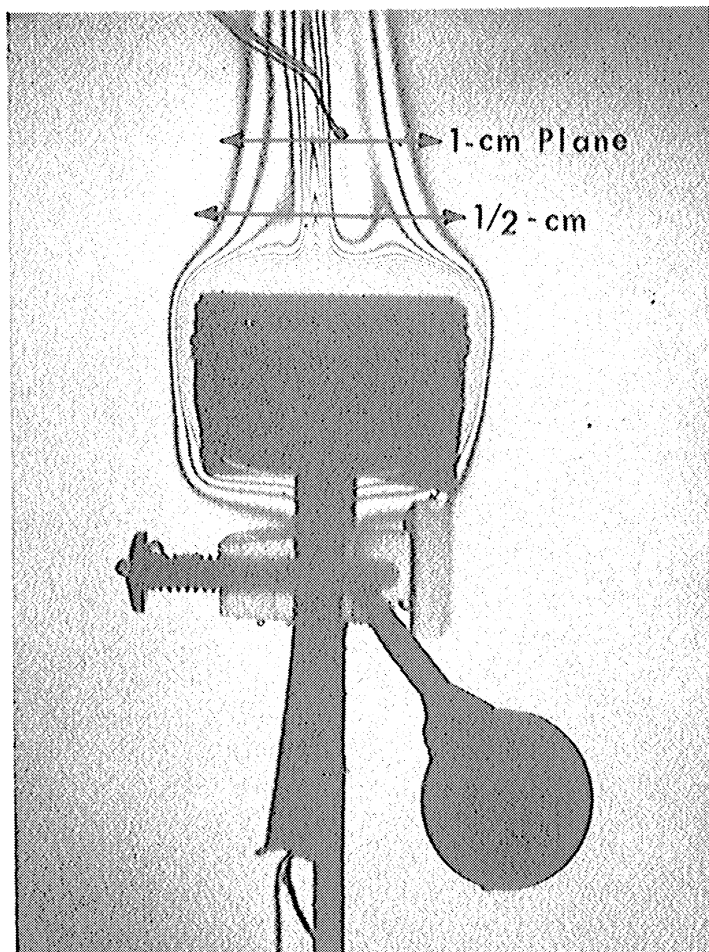


Fig. 5-8. Typical View of the Interferogram of Heated Plate in Water

Figures 5-9 and 5-10 show the reconstructed temperature field at an elevation of 1 cm above the heated surface. Figure 5-9 is an isometric plot of the temperature field in this plane. Figure 5-10 is an isothermal contour plot of the temperature field. These results were obtained using the sample method with three resolution elements, each in the x and y direction, i.e., $M = 3$, $N = 3$.

The thermocouple probe, which was inserted into the flow to obtain an independent temperature measurement, indicated a temperature

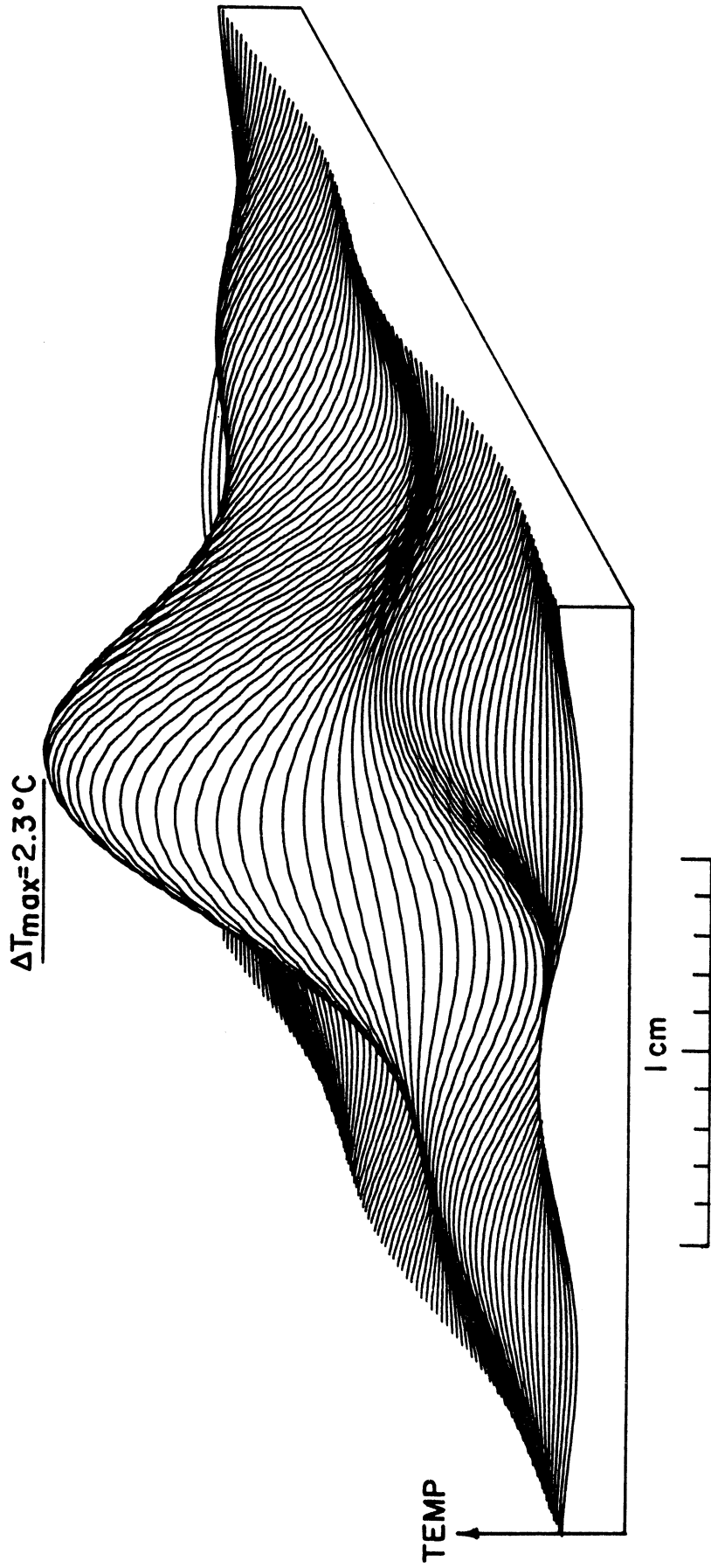


Fig. 5-9. Isometric Plot of the Reconstructed Temperature Field in a Plane 1 cm Above the Heated Plate (Sample Method, $M = 3$, $N = 3$)

LABEL	ΔT (deg. C)
1	.3
2	.6
3	1.1
4	1.6
5	2.1

$$\Delta T_{tc} = 1.9$$

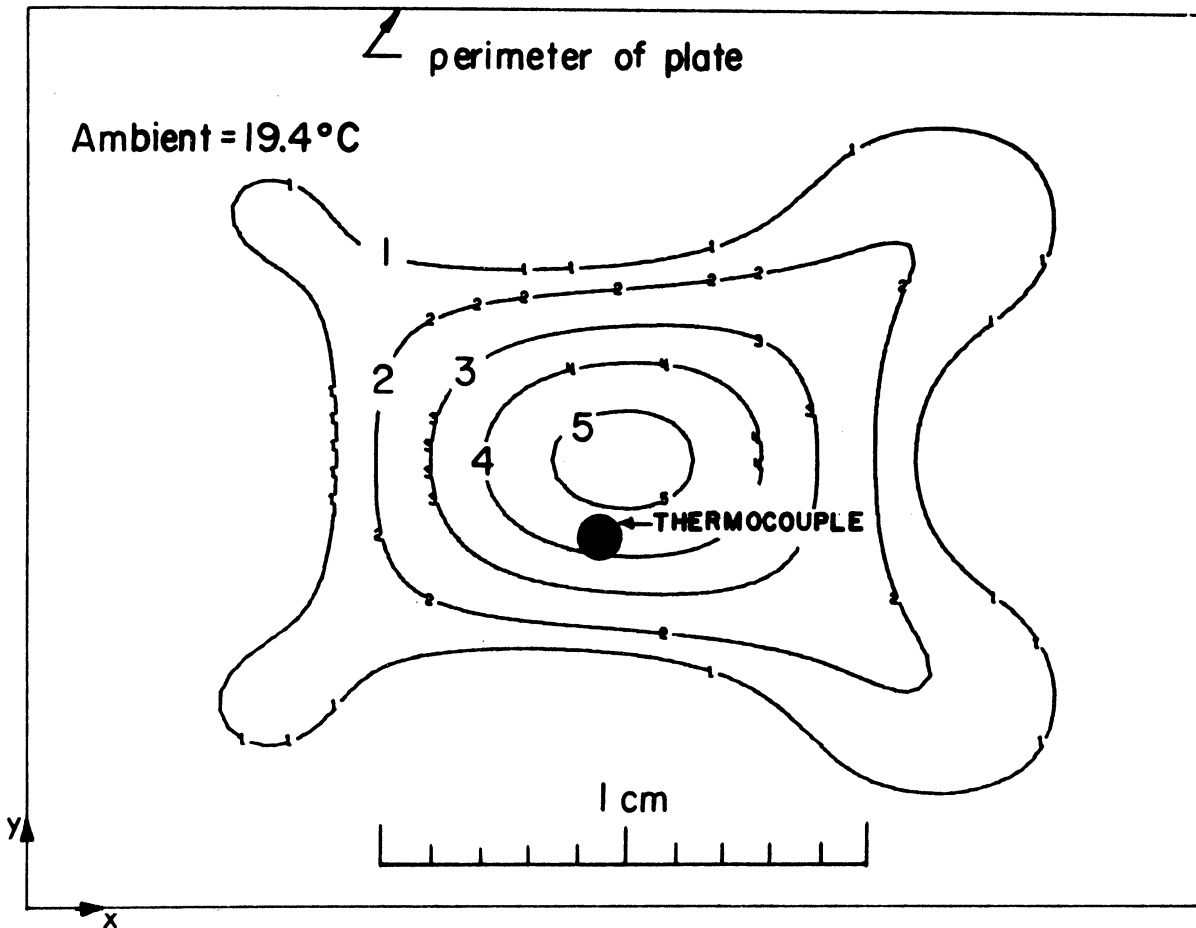


Fig. 5-10. Isothermal Contour Plot for a
Plane 1 cm Above the Heated Surface
(Sample Method, $M = 3$, $N = 3$)

rise of 1.9°C above ambient. The thermocouple, as shown in Fig. 5-10 occupies a region between the isothermal contour lines of 1.6 and 2.0°C . Once again the thermocouple occupies a substantial region in the convective plume. The situation is clearly better, however, than that in the heated wire experiment. A conservative estimate of the difference between the two measurements is about 10%. The asymmetry in the reconstructed image field about the bisector of the plate in the y direction is due to either of two effects. The asymmetry could be a real effect caused by the wandering of the plume, or it could be due to errors in the reconstruction. Since the lower plane is more symmetric, the former is believed to be the case.

Figure 5-11 is an isothermal contour plot showing the temperature field on the plane located 0.5 cm above the heated surface. The contour plot was obtained using the sample method with four resolution elements in the y direction and three elements in the x direction, i.e., $M = 3$, $N = 4$.

A qualitative study of the shape of the convective plume above horizontal heated surfaces has been made by Husar and Sparrow [54]. They used flow-visualization techniques to detect the fluid motions around heated horizontal surfaces. In their experimental study, fluid motions were made visible by local changes in the color of the fluid itself. The color changes result from slight changes in the pH of the fluid (cf. [55]).

LABEL	ΔT (deg.C)
-------	--------------------

1	.3
2	.6
3	1.1
4	1.6
5	2.1
6	2.6
7	3.1

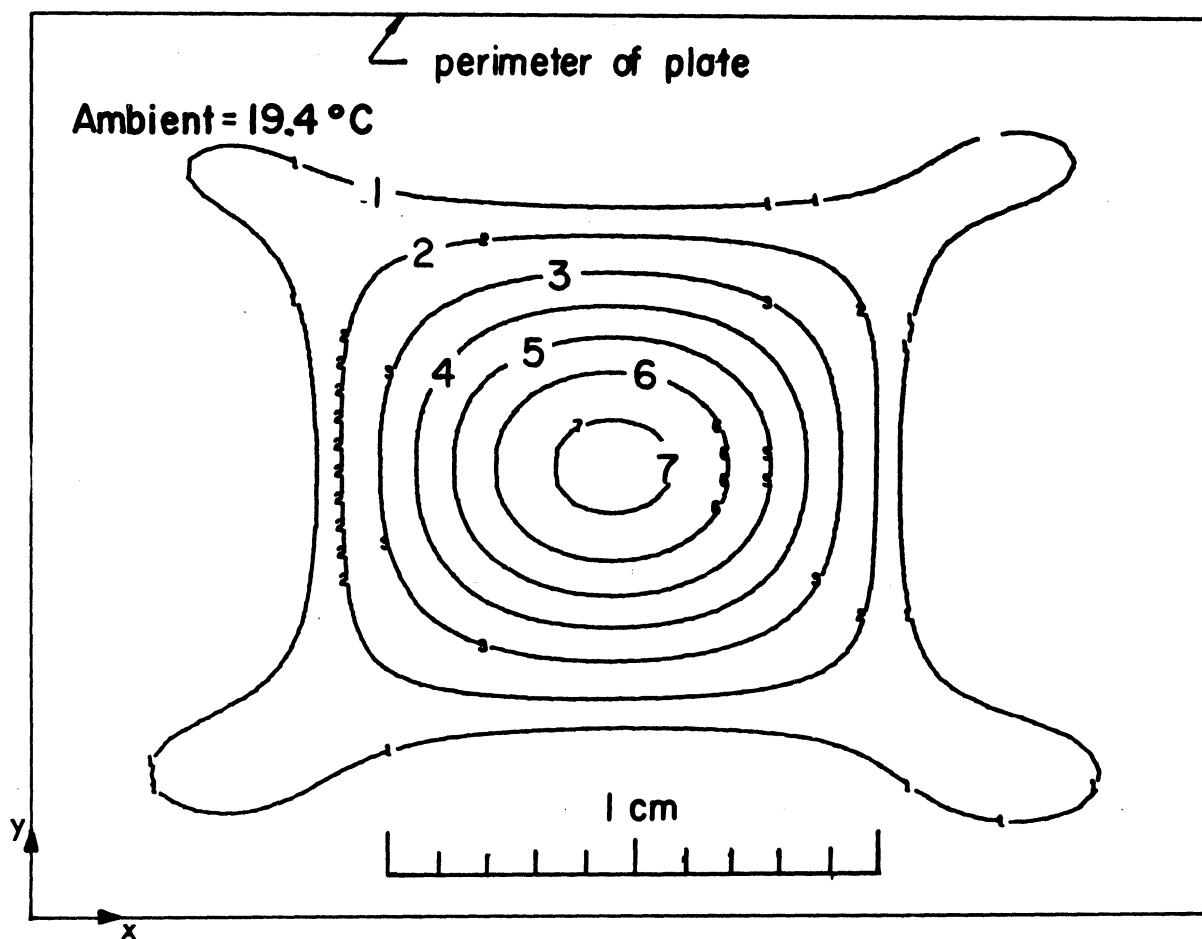


Fig. 5-11. Isothermal Contour Plot for a Plane 0.5 cm Above the Heated Plate (Sample Method, $M = 3$, $N = 4$)

Figure 5-12 is a sketch of two of their experimental results. Figure 5-12a is a view from directly above a heated rectangular plate in water; Fig. 5-12b is a view from directly above a heated square plate in water. The dark lines on the plate indicate regions of stagnation on the surface of the plate. The results of this and other experiments lead Husar and Sparrow [54] to argue that a basic characteristic of free convective flow adjacent to platforms with corners is the partitioning of the flow field. There is no fluid flow across a partition line. Each partition line is a central element of a vertically ascending buoyant plume. The partition lines are created by the intersection of the flows streaming inward along the plate surface from the separate edges. The reconstructed temperature fields shown in Figs. 5-11 and 5-12 are in agreement with the partitioning effects proposed by Husar and Sparrow [54]. The regions of high velocity along the partition lines must have greater buoyancy forces. Since the buoyancy forces are directly related to the temperature, these regions must also have higher temperatures. Figures 5-10 and 5-11 show that the temperature in these regions is indeed higher.

The temperature field above the heated plate was also reconstructed using the grid and frequency plane restoration methods. Figures 5-13 to 5-15 show several reconstructions using the other techniques. The results are in good agreement with those obtained with the sample method.

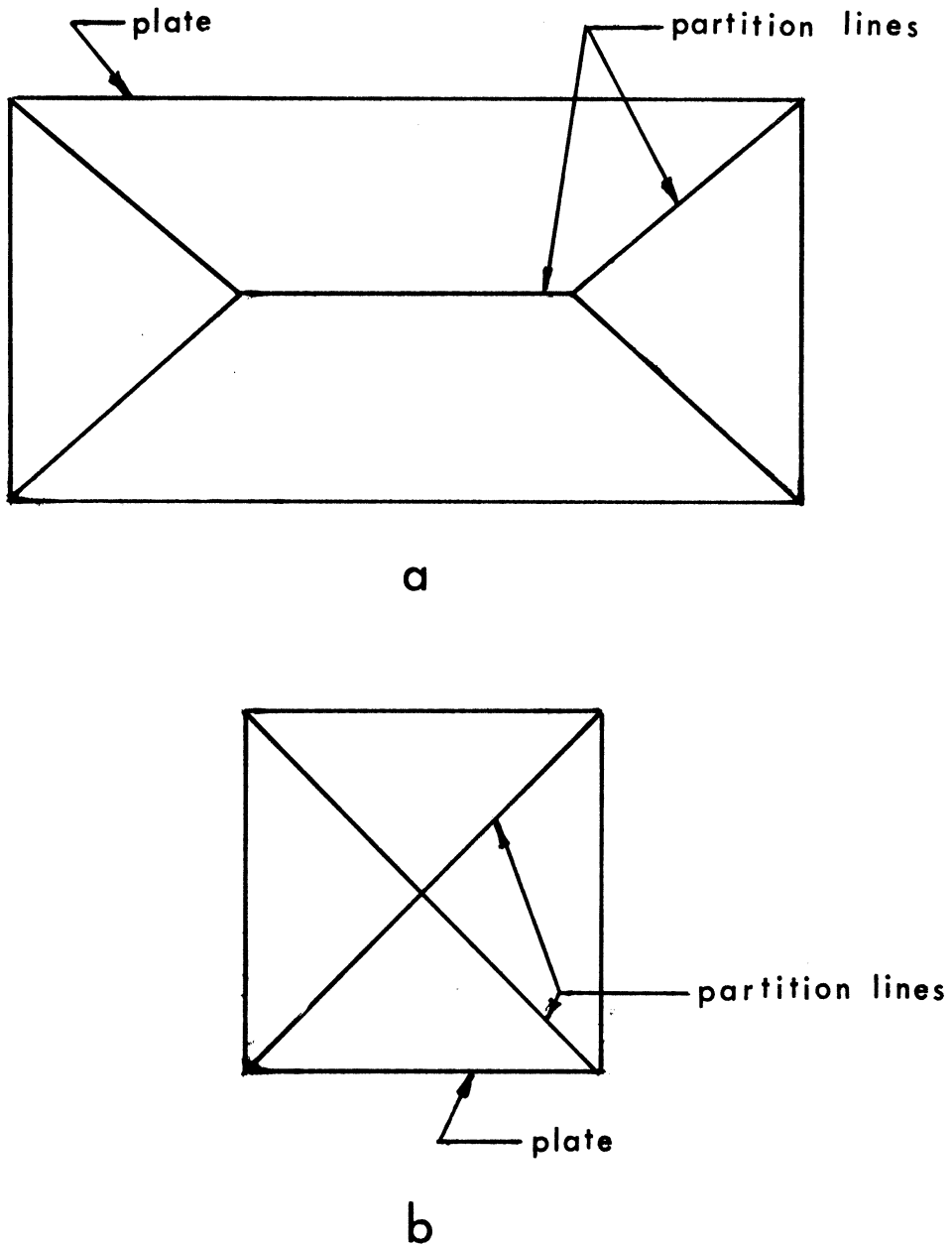


Fig. 5-12. Sketch of Results of Flow Visualization
Obtained by Husar and Sparrow [54] for Free
Convection Above Horizontal Heated Surface

- (a) Rectangular Plate with the Long Side Twice
the Short Side
- (b) Square Plate

LABEL	ΔT (deg C)
1	.3
2	.6
3	1.1
4	1.6
5	2.1

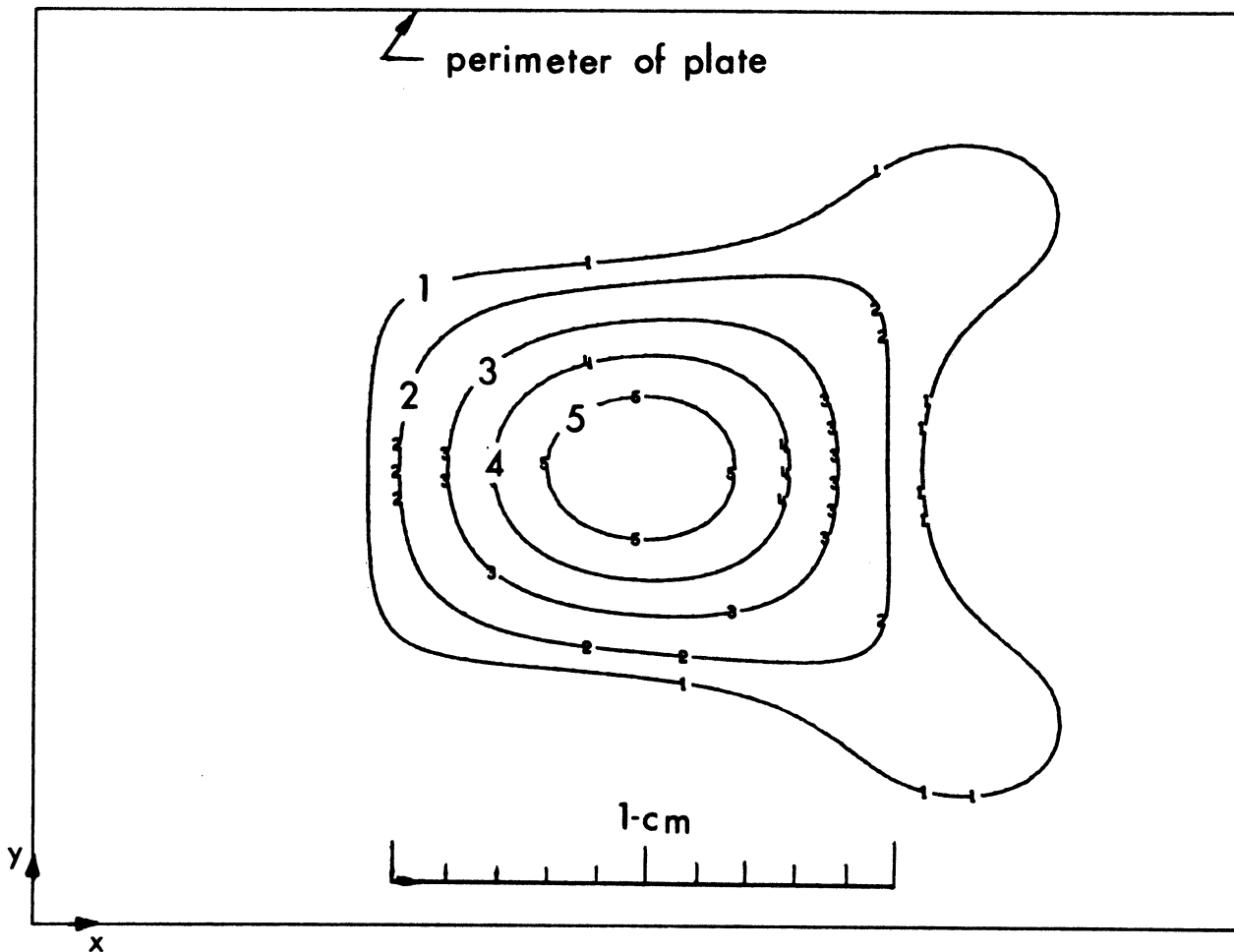


Fig. 5-13. Isothermal Contour Plot for Plane 1 cm Above Heated Surface (Grid Method, $M = 3$, $N = 3$)

LABEL	ΔT (deg C)
1	.3
2	.6
3	1.1
4	1.6
5	2.1
6	2.6
7	3.1

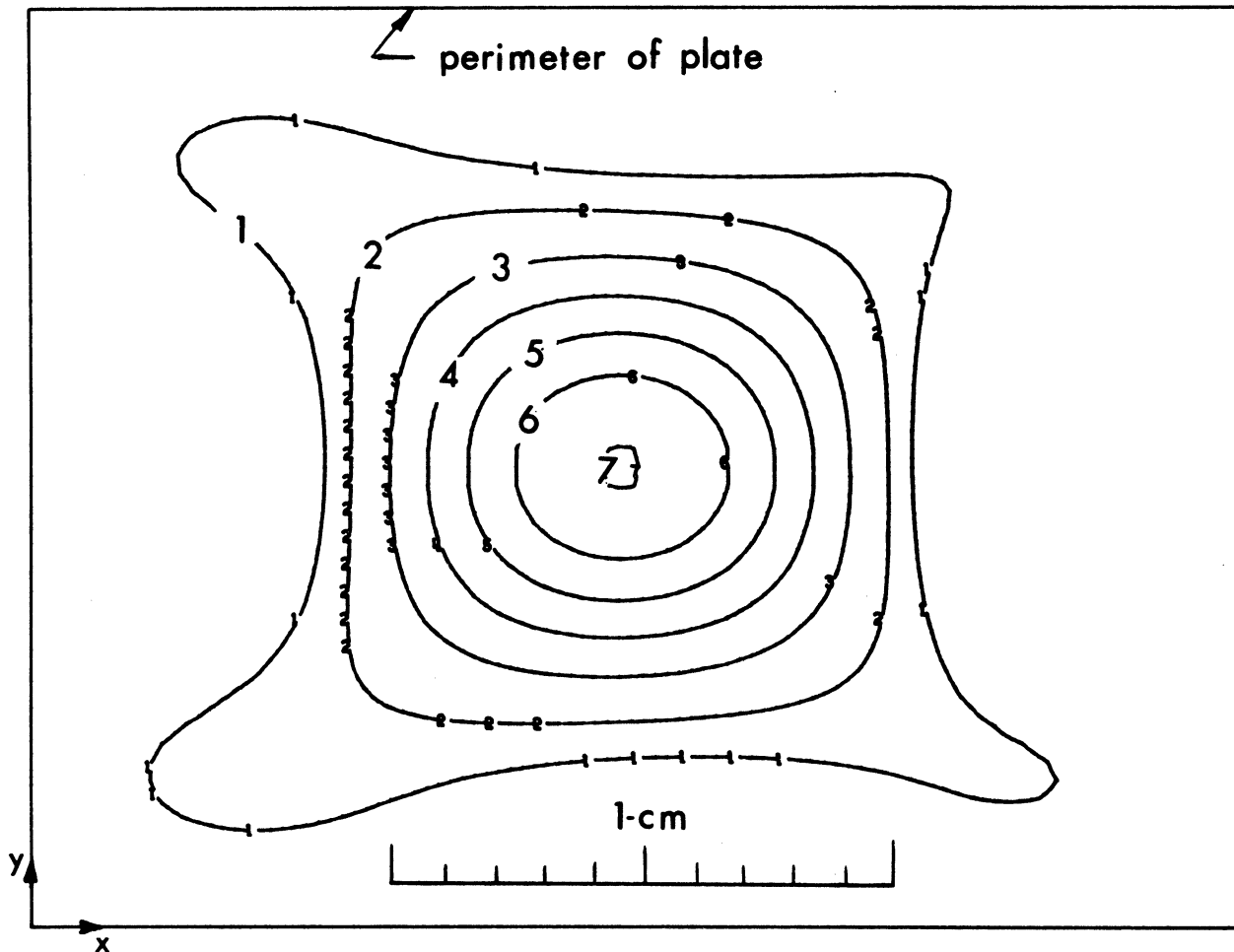


Fig. 5-14. Isothermal Contour Plot for Plane 0.5 cm Above Heated Surface (Grid Method, $M = 3$, $N = 4$)

LABEL	ΔT (deg C)
1	.3
2	.6
3	1.1
4	1.6
5	2.1
6 <td 2.6	
7	3.1

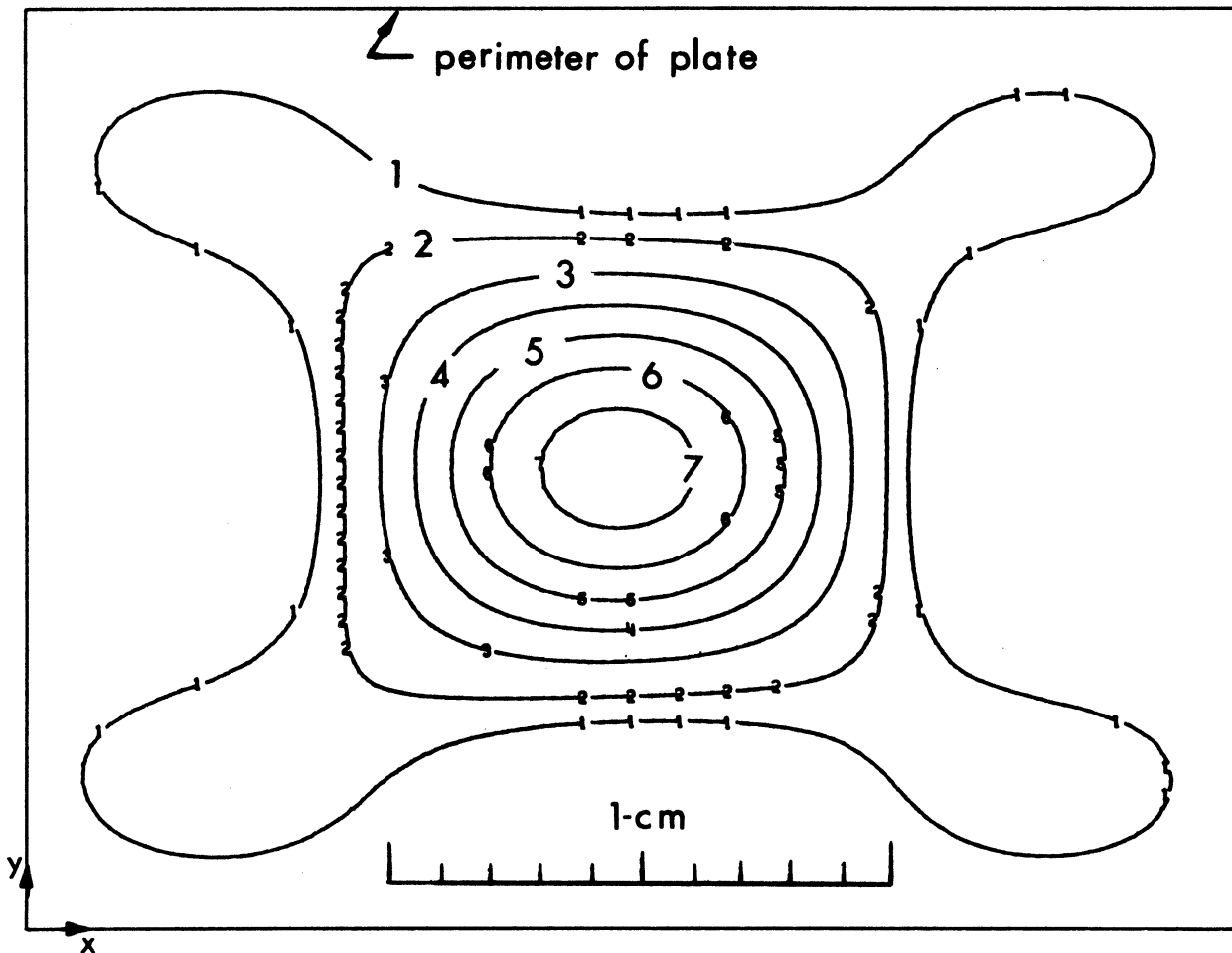


Fig. 5-15. Isothermal Contour Plot for Plane 0.5 cm Above Heated Surface (Frequency Plane Restoration, $M = 3$, $N = 4$)

CHAPTER 6

SUMMARY AND CONCLUSIONS

Reconstruction by Fourier Analysis and Series Expansions

The integral equation which describes the relationship between a refractive index field and the optical pathlength in the refractionless limit can be written as

$$\int_S f(x, y) dS = \phi(\rho, \theta) \quad (6-1)$$

The unknown functions $f(x, y)$ is the refractive index relative to a known reference. The pathlength function $\phi(\rho, \theta)$ is the value of the integral of the index field along straight lines (see Fig. 2-2). This is the data obtained by interferometric experiments. Fourier analysis of the pathlength data provides the basic relationship

$$F(R, \theta + \pi/2) = \int_{-\infty}^{+\infty} \phi(\rho, \theta) \exp(-i2\pi R\rho) d\rho \quad (6-2)$$

where $F(R, \theta + \pi/2)$ is the two-dimensional Fourier transform of $f(x, y)$ and the notation is shown in Fig. 2-2. This shows that for a fixed value of θ , the one-dimensional Fourier transform of the pathlength function can be used to determine a single line through the origin in the Fourier transform of $f(x, y)$.

Equation (6-2) can be used to obtain a direct relationship between the pathlength function and the relative refractive index field.

The results of the analysis of Chapter 2 show that (see Fig. 2-2)

$$f(r, \phi) = \frac{1}{2\pi^2} \int_{-\pi/2}^{+\pi/2} d\theta \int_{-\infty}^{+\infty} \frac{\partial \Phi(\rho, \theta) / \partial \rho}{r \sin(\phi - \theta) - \rho} d\rho \quad (6-3)$$

Both Eqs. (6-2) and (6-3) provide exact reconstructions of the index field if the pathlength function $\phi(\rho, \theta)$ is a continuous function of ρ and θ for all θ in $[0, \pi]$. In real situations, however, the pathlength function is only known for sampled values of ρ and θ . The use of discrete data produces a reconstruction which is an approximation to $f(x, y)$. Exact recovery of the index field is not possible using a finite number of samples because the bandwidth of the spatially limited refractive index field is infinite. The sampling theorem states that for complete recovery of a spatially limited function, the samples must be taken at an infinite rate. The approximation of $f(x, y)$ will be acceptable, however, if the samples are taken at a rate equal to twice an effective bandwidth of $f(x, y)$. If samples are taken at a rate less than this, aliasing can introduce large errors. If samples are not available at a rate equal to twice the effective bandwidth of $f(x, y)$, then a limited resolution reconstruction can be found with a bandwidth B_1 given by

$$B_1 = 2B_S - B \quad (6-4)$$

where B is the effective bandwidth of $f(x, y)$ and $(1/2B_S)$ is the pathlength sample rate available.

Theoretically, only as many independent pathlength samples need to be collected as there are degrees of freedom in the reconstructed index

field. In practice, however, the sensitivity of the reconstruction process to input errors makes it necessary to introduce redundant data to obtain acceptable reconstructions. The effect of redundant data is to average over the redundant measurements. If the errors of each pathlength sample are uncorrelated, the effect of errors can be suppressed at a rate proportional to the square root of the number of pathlength samples.

The analytical solution requires that data be collected over a full 180° angle of view. Since it is particularly difficult to obtain a 180° angle of view in a holographic system, several reconstruction procedures which use series expansions for $f(x, y)$ were developed. The simplest series expansion represents the refractive index field as a number of rectangular grid elements; within each element, the refractive index is a constant. A ray through the index field generates a linear algebraic equation where the variables are the values of refractive index within each cell. The overdetermined and inconsistent set of equations generated by collecting pathlength data for many more rays than unknowns is solved for the least-mean-square solution using Golub's [42] computational procedure.

It is shown that the series used in the Whittaker-Shannon sampling theorem can be directly integrated to generate a set of simultaneous equations which can be solved for the sample values. In this procedure, referred to as the sample method, a set of equations, which is considerably more consistent than the equations

of the grid method, is generated. This forms the basis for a better reconstruction process.

Another approach is to apply the techniques developed by J. L. Harris [16] in the area of image resolution; the sampling theorem is applied in the frequency plane. In this way a system of linear equations is generated and solved to determine the coefficients of the Fourier series of $f(x, y)$. The Fourier series determined by this procedure provides a bandlimited (i.e., limited resolution) representation of the refractive index field.

The experimental application of the various reconstruction techniques is demonstrated by using them to determine the temperature field in a steady-state convective plume above a short heated wire. The experimental system uses phase grating illumination to eliminate the effects of coherent speckle. If phase grating illumination is not used, the speckle pattern will prevent the fringe pattern to be resolved.

The techniques are also applied to determine the temperature profile above a horizontal plate heated electrically in water. The reconstructed temperature field reveals the same basic characteristics that have been noted by Husar and Sparrow [54] using flow visualization techniques. The temperature measured by a thermocouple in the plume and the reconstructed temperature are in good agreement.

Conclusions

It has been shown that the pathlength data sampled at discrete locations and over a limited viewing angle can be used to determine a

limited-resolution reconstruction of an asymmetric refractive index field. The reconstructed refractive index field can then be used to determine the temperature field.

The index field can be reconstructed using a number of different procedures. The techniques which can be used with less than a 180° angle of view all require solving a system of simultaneous equations. The effect of an angle of view of less than 180° is to increase the sensitivity of the reconstruction process to errors. These errors are introduced by using a truncated series expansion to represent a function with an infinite number of degrees of freedom. The effect of input errors on the reconstruction can be reduced by using greater redundancy and/or by decreasing the number of resolution cells that lie along the ray in the mean-viewing direction.

When a complete 180° viewing angle is available, both the analytical solution and the series solutions provide results of comparable accuracy. Since the analytical techniques only involve fast Fourier transform algorithms, rather than the matrix operations required by the series solutions, the analytical solutions result in a considerable savings in computer time.

REFERENCES

1. W. Hauf and U. Grigull, Advances in Heat Transfer, 6, Academic Press, New York, 1970.
2. M. Francon, Optical Interferometry, Academic Press, New York, 1966.
3. W. H. Steel, Interferometry, Cambridge University Press, London, 1967.
4. R. J. Collier, et al., Optical Holography, Academic Press, New York, 1971.
5. M. Born and E. Wolf, Principles of Optics, 3rd Ed., Pergamon Press, Oxford, 1965.
6. R. N. Bracewell, Aust. J. of Phys., 9, 198 (1956).
7. R. N. Bracewell and A. C. Riddle, Astrophys. J., 150, 427 (1967).
8. J. H. Taylor, Astrophys. J., 150, 421 (1967).
9. P. A. O'Brien, Mon Not. Roy. Astr. Soc., 113, 597 (1953).
10. S. F. Smerd and J. P. Wild, Philo. Mag., Series 8, 2, 119 (1957)
11. D. J. DeRosien and A. Klug, Nature, 217, 130 (1968).
12. A. Klug, Phil. Trans. Roy. Soc. Lond., B261, 173 (1971).
13. R. A. Crowther, et al., Proc. Roy. Soc., London, A317, 319 (1970).
14. R. A. Crowther, Phil. Trans. Roy. Soc. Lond., B261, 221 (1971).
15. W. Hoppe, Optik, 29, 617 (1969).
16. J. Harris, J. Opt. Soc. Am., 54, 931 (1964).
17. O. J. Tretiak, et al., 8th International Conference on Med. and Biol. Engineering, Chicago, Illinois, July 20-25, 1969.
18. O. J. Tretiak, et al., Digital Proc. Conf., Columbia, Missouri, October 6-8, 1971.
19. P. Rowley, J. Opt. Soc. Am., 59, 1496 (1969).
20. M. Berry and D. Gibbs, Proc. Roy. Soc. Lond., A314, 143 (1970).

21. L. H. Tanner, J. Sci Instrum., 3, 987 (1970).
22. W. Alwang, et al., Item 1, Final Report, Pratt and Whitney, PWA-3942 [NAVAIR (Air-602) Contract], 1970.
23. H. G. Junginger and W. van Haeringer, Opt. Comm., 5, 1 (1972).
24. G. Ramachandran, Proc. Indian Acad. Sci., LXXIV, 14 (1971).
25. G. Ramachandran and A. Lakshminarayanan, Proc. Natl. Acad. Sci., U.S., 68, 2236 (1971).
26. C. D. Maldonado, J. of Math Phys., 6, 1935 (1965).
27. A. Bhatia and E. Wolf, Proc. Cambridge Phil. Soc., 50, 40 (1954).
28. C. D. Maldonado and H. Olsen, J. Opt. Soc. Am., 56, 1305 (1966).
29. R. D. Matulka and D. J. Collins, J. Appl. Phys., 12, 1109 (1971).
30. R. D. Matulka, Ph.D. Thesis, Naval Postgraduate School (1970).
31. R. Gordon, et al., J. Theor. Biol., 29, 471 (1970).
32. R. Bender, et al., J. Theor. Biol., 29, 483 (1970).
33. G. T. Herman and S. Rowland, J. Theor. Biol., 33, 213 (1971).
34. R. N. Bracewell, The Fourier Transfer and Its Applications, McGraw Hill, New York, 1965.
35. A Papoulis, Systems and Transforms with Applications in Optics, McGraw-Hill, New York, 1968.
36. W. T. Cochran, et al., Proc. IEEE, 55, 1664 (1967).
37. R. V. Churchill, Fourier Series and Boundary Value Problems, McGraw Hill, New York, 1941.
38. J. W. Cooley and J. W. Tukey, Math. Comp., 19, 297 (1965).
39. D. A. Linden, Proc. IRE, 47, 1219 (1959).
40. A. Papoulis, Proc. IEEE, 55, 1677 (1967).
41. F. Hildebrand, Introduction to Numerical Analysis, McGraw-Hill, New York, 1956.
42. G. Golub, Numer. Math., 7, 206 (1965).
43. J. H. Wilkinson and C. Reinsch, Handbook for Automatic Computation-Linear Algebra, Vol. 2, Springer-Verlag, Berlin, 1971.

44. W. A. Brown, Trans. IRE, IT7, 269 (1961).
45. G. Forsythe and C. Moler, Computer Solution of Linear Algebraic Systems, Prentice Hall, Englewood Cliffs, New Jersey, 1967.
46. B. R. Frieden, J. Opt. Soc. Am., 62, 511 (1972).
47. J. B. Traub, Iterative Methods for the Solution of Equations, Prentice Hall, Englewood Cliffs, New Jersey, 1964.
48. H. Eisenberg, J. of Chem. Eng., 43, 3887 (1965).
49. L. H. Tanner, J. Sci. Instrum., 44, 1011 (1967).
50. L. H. Tanner, J. Sci. Instrum., Sec. 2, 1, 517 (1968).
51. L. I. Goldfisher, J. Opt. Soc. Am., 55, 247 (1965).
52. C. M. Vest and D. W. Sweeney, Appl. Opt., 9, 2321 (1970).
53. G. D. Currie, Appl. Opt., 8, 1068 (1969).
54. R. B. Husar and E. M. Sparrow, Int. J. Heat Mass Transfer, 11, 1206 (1968).
55. D. J. Baker, J. Fluid Mech., 26, 573 (1966).