RECONSTRUCTION OF RADIATING SOURCES

A thesis presented for the degree of
Doctor of Philosophy in Electrical Engineering
in the University of Canterbury,
Christchurch, New Zealand.

by

P.J. NAPIER B.E.(Hons.)

1971
ABSTRACT

The problem of obtaining phase information when radiating source distributions are reconstructed by Fourier transforming their measured radiated field patterns is examined. The two cases examined in detail are those of determining an antenna aperture distribution from its far-field radiation pattern and of reconstructing an astronomical brightness temperature distribution from data obtained using an interferometer.

New techniques are presented which allow the aperture distribution of an antenna to be reconstructed from only the modulus of a far-field radiation pattern. The techniques, which are based on the principle of holography, are described theoretically and are shown to be practical using the results of experiments performed with acoustic antennas. The methods are also shown to be applicable in the near-field of an antenna.

Using the concept of the complex zeros of an interferogram, it is shown how phase errors can be identified in the data obtained from astronomical interferometers. In those cases in which the measured intensity is more accurate than the measured phase (it is shown how to process the measured data to determine this) it is shown how to compute an improved phase from the measured intensity. The problem of reconstructing brightness-temperature distributions from intensity only interferometer measurements is also discussed. The results obtained by processing actual measured data using the new techniques demonstrates their usefulness.
ACKNOWLEDGEMENTS

I am especially grateful to my supervisor Mr R.H.T. Bates for his guidance and encouragement. His assistance over the duration of the project has been invaluable.

I also thank my wife Lynette for her patience and encouragement.

The financial assistance of the University Grants Committee and the New Zealand Post Office is gratefully acknowledged.
# TABLE OF CONTENTS

Abstract 
Acknowledgements 
Glossary vii 
Preface x

CHAPTER 1: Antenna Aperture Distributions and Radiation Pattern Phase 1 
1.1 Identification of Errors in Antennas 1 
1.2 The Aperture Distribution and Radiation Pattern of an Antenna 3 
1.2.1 The Fourier Transform Relationship 3 
1.2.2 Determination of the Aperture Distribution 6 
1.3 Measurement of Radiation Pattern Phase 8 
1.4 Radiation Pattern Measurement in Near Field 10 
1.5 Previous Users of Modulus Only Pattern Information 11

CHAPTER 2: A Holographic Approach to Radiation Pattern Measurement - Theory 15 
2.1 Introduction 15 
2.2 Non Optical Holography 16 
2.2.1 Microwave Holography 16 
2.2.2 Acoustic Holography 19 
2.3 A Holographic Approach to Antenna Radiation Pattern Measurement 20 
2.3.1 Basic Theory 20 
2.3.2 H.P.M. Techniques With Partial Separation 26
2.3.3 H.P.M. Techniques With Point Source Reference 28
2.3.4 Interpretation of H.P.M. Techniques 29
2.4 Accuracy of H.P.M. Techniques 30
   2.4.1 A Criterion for Estimating the Accuracy of the H.P.M. Techniques 30
   2.4.2 The Effect of Pattern Truncation 32
   2.4.3 The Effect of Measurement Noise 35
   2.4.4 A Practical Error Estimate 37

CHAPTER 3: A Holographic Approach to Radiation Pattern Measurement - Experiment 43
   3.1 Experimental Arrangement 43
   3.2 Application of Basic H.P.M. Technique With Complete Separation and Arbitrary Reference Source 45
      3.2.1 Procedure and Results of Experiment 45
      3.2.2 Discussion of Basic H.P.M. Technique 47
   3.3 Application of H.P.M. Techniques to the Measurement of Array Aperture Distributions 50
      3.3.1 The Calibration of an Array 50
      3.3.2 Application of H.P.M. With Complete Separation 51
      3.3.3 Application of H.P.M. with Partial Separation 55
      3.3.4 Discussion of Array Measurement Techniques 57
   3.4 H.P.M. Techniques With a Point Source Reference 60
      3.4.1 Provision of a Point Reference Source Using a Fixed Reference Antenna 60
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.2</td>
<td>H.P.M. Technique With Point Source Reference and Complete Separation</td>
<td>61</td>
</tr>
<tr>
<td>3.4.3</td>
<td>H.P.M. Technique With Point Source Reference and Partial Separation</td>
<td>64</td>
</tr>
<tr>
<td>3.4.4</td>
<td>H.P.M. Technique With Off Axis Point Source Reference</td>
<td>65</td>
</tr>
<tr>
<td>3.4.5</td>
<td>Discussion of H.P.M. Techniques with Point Source References</td>
<td>69</td>
</tr>
<tr>
<td>CHAPTER 4:</td>
<td>Extensions to Holographic Pattern Measurement Theory</td>
<td>83</td>
</tr>
<tr>
<td>4.1</td>
<td>Use of a Reference Signal With Phase Shift Linear With Angle</td>
<td>83</td>
</tr>
<tr>
<td>4.2</td>
<td>A H.P.M. Technique Analogous to Gabor In-Line Holography</td>
<td>88</td>
</tr>
<tr>
<td>4.3</td>
<td>Use of H.P.M. Techniques in the Near Field</td>
<td>90</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Measurement of Narrow Beam Antennas</td>
<td>90</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Phase Correction of the Reference Signal</td>
<td>96</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Angular Fourier Series</td>
<td>97</td>
</tr>
<tr>
<td>4.4</td>
<td>Advantage of H.P.M. Techniques in a Distorting Medium</td>
<td>103</td>
</tr>
<tr>
<td>CHAPTER 5:</td>
<td>The Zeros of Entire Functions</td>
<td>112</td>
</tr>
<tr>
<td>5.1</td>
<td>Entire Function Theory</td>
<td>112</td>
</tr>
<tr>
<td>5.1.1</td>
<td>Fourier Transforms in the Complex Plane</td>
<td>112</td>
</tr>
<tr>
<td>5.1.2</td>
<td>Complex Zeros and the Sampling Theorem</td>
<td>115</td>
</tr>
<tr>
<td>5.1.3</td>
<td>The Source Distributions Defined by $</td>
<td>F(u)</td>
</tr>
</tbody>
</table>
5.2 Review of Literature on Applications of Complex Zero Theory
5.3 Some Useful Theorems on Zeros of Entire Functions of Exponential Type

CHAPTER 6: Application of Complex Zero Techniques to Astronomical Interferometry
6.1 Astronomical Interferometry
   6.1.1 The Fourier Transform Relationship in Radio Interferometry
   6.1.2 The Problem of Phase Measurement in Radio Astronomy
   6.1.3 Interferometry in Optical Astronomy
6.2 Identification of Errors in a Measured Interferogram
   6.2.1 Preliminary Definitions
   6.2.2 Three Examples
   6.2.3 Comparison of Extents of Distributions
   6.2.4 Negative Parts in the Brightness Temperature
6.3 Removal of Phase Errors in Interferometry
   6.3.1 Complex Zeros of the Interferogram
   6.3.2 Interferogram Phase Correction Procedure
   6.3.3 Discussion of Phase Correction Procedure
6.4 The Reconstruction of Brightness Temperatures
   From Intensity Only Interferograms
   6.4.1 Intensity Only Interferograms
   6.4.2 Three Examples
6.4.3 Intensity Interferogram Reconstruction Procedure 153
6.4.4 Discussion of Intensity Only Reconstruction 158

CHAPTER 7: Combination of Holographic and Complex Zero Reconstruction Techniques 180
7.1 Introduction 180
7.2 Use of an A Priori Estimate of the Source Distribution 181
  7.2.1 Description of the Problem 181
  7.2.2 A Reconstruction Procedure Using an A Priori Estimate of a Source Distribution and the Measured Modulus of its F.T. 182
  7.2.3 An Example of the Use of Procedure 186
7.3 Application of the Holographic Principle to Interferometry 189
7.4 Deconvolution Using Complex Zeros 190
  7.4.1 Statement of the Problem 190
  7.4.2 Review of Methods of Deconvolution 191
  7.4.3 Deconvolution Without Phase Information 193
  7.4.4 Applications of the Deconvolution Procedure 196
7.5 Reconstruction of Double, Separated Sources 198

CHAPTER 8: Computational Considerations 210
8.1 Computation of Fourier Transforms 210
8.2 Computation of Zeros 214
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.2.1 Computation of Zeros Using Polynomial Method</td>
<td>214</td>
</tr>
<tr>
<td>8.2.1.1 Zeros of F(w)</td>
<td>215</td>
</tr>
<tr>
<td>8.2.1.2 Theory of the Method for Zeros of R(w)</td>
<td>215</td>
</tr>
<tr>
<td>8.2.1.3 Accuracy of the Polynomial Zero Computation Method</td>
<td>217</td>
</tr>
<tr>
<td>8.2.1.4 Zero Refinement Procedure</td>
<td>219</td>
</tr>
<tr>
<td>8.2.2 Other Zero Computation Methods</td>
<td>222</td>
</tr>
<tr>
<td>8.3 Reconstruction of F(u) and f(x) From the Zeros of F(w)</td>
<td>226</td>
</tr>
<tr>
<td>CHAPTER 9: Conclusion</td>
<td>231</td>
</tr>
<tr>
<td>9.1 Suggestions for Future Research on H.P.M. Techniques</td>
<td>231</td>
</tr>
<tr>
<td>9.2 Suggestions for Future Research on Complex Zero Techniques</td>
<td>232</td>
</tr>
<tr>
<td>Appendices</td>
<td>235</td>
</tr>
<tr>
<td>References</td>
<td>245</td>
</tr>
</tbody>
</table>
GLOSSARY
Symbols, Abbreviations and Definitions

Unless otherwise defined, symbols have the meanings given below.

arg : argument. e.g. F(u) = |F(u)| exp(arg F(u)).
D : separation between g(x) and h(x).
\text{erf}^{-1}(x): \text{inverse error function}
\[ z = \text{erf}^{-1}(x) \text{ means } x = \frac{2}{\sqrt{\pi}} \int_{0}^{z} \exp(-y^2) dy \]
e(x) : correlation of g(x) and h(x).
\text{E}(u) : \text{F.T. of e(x)}.
f(x), g(x), h(x): \text{radiating source distributions; aperture distributions for antennas or brightness temperature distributions for astronomical sources.}
\text{F}(u), \text{G}(u), \text{H}(u): \text{F.T.s of source distributions; radiation patterns for antennas or interferograms of astronomical sources.}

\text{F.T.} : \text{Fourier Transform.}
\text{F.F.T.} : \text{Fast Fourier Transform.}
\text{H.P.M.} : \text{Holographic pattern measurement.}
j : \text{square root of } -1, \quad j^2 = -1.
k : \text{wavenumber } \quad k = \frac{2\pi}{\lambda}.
L, L_1, L_2 : \text{widths of } f(x), g(x), h(x) \text{ respectively.}
L_0 : \text{the greater of } L_1 \text{ or } L_2.
N(u) : \text{measurement noise in } |F(u)|.
\text{P}(w) : \text{a factor of } F(w).
\text{Q}(w) : \text{a factor of } F(w).
\text{rect}(x) : \text{rectangular gate function, } \text{rect}(x) = 1, \quad |x| < \frac{1}{2}
\quad = 0, \quad |x| > \frac{1}{2}.
r(x) : \text{F.T. of } R(u), \text{ autocorrelation of } f(x).
R : separation between test and measuring antennas for radiation pattern measurement.
R.F. : radio frequency.
R.H.S. : right hand side.
R(u) : Intensity of F(u) \( R(u) = |F(u)|^2 \).
sinc (x): sampling function, sinc(x) = \(\frac{\sin(x)}{x}\).
S : strength of reference signal.
tri(x) : triangular gate function, \(\text{tri}(x) = 1 - 2|\ x|, x < \frac{1}{2}\)
\( = 0, \quad x > \frac{1}{2}\)
\( T_T, T_R \) : positions on turntable of test and reference antennas.
T(w) : a factor that flips zeros of F(w).
u : real F.T. variable; angular variable for radiation patterns or baseline variable for interferograms.
\( u_{\text{max}} \) : maximum value of u available from a measurement.
U.H.P. : Upper half plane.
v : imaginary F.T. variable.
V : a noise threshold.
w : complex F.T. variable, \( w = u + jv \).
\( w_n, w_{gn}, w_{hn} \) : zeros of \( F(w) \), \( G(w) \), \( H(w) \) respectively.
x : source coordinate; aperture coordinate for antennas or angular coordinate for astronomical sources.
\( \beta \) : half width of f(x).
\( \delta(x) \) : Dirac delta function.
\( \Delta u \) : sample spacing on the u axis.
\( \zeta_n \) : zero of \( R(w) \) in U.H.P. \( \zeta_n = p_n + jq_n \).
\( \theta \) : angle measured from direction normal to aperture plane.
\( \lambda \) : wavelength.
\( \omega \) : radian frequency.
\( \psi(u) \) : Phase error in F(u).
\[ \prod_{m=1}^{M} (n) \text{ : Continued product for } m \text{ from 1 to } M, \text{ with } n^{th} \text{ term missed out.} \]

\[ \{ \} \text{ : notation for a set.} \]

\[ U \text{ : union.} \]

\[ \cap \text{ : intersection.} \]

\[ * \text{ : complex conjugate, } w^* = u - jv. \]

\[ * \text{ : convolution, } g(x) * h(x) = \int_{-\infty}^{\infty} g(\alpha)h(x-\alpha)d\alpha. \]

\[ \otimes \text{ : correlation, } g(x) \otimes h(x) = \int_{-\infty}^{\infty} g(\alpha)h(\alpha+x)d\alpha. \]

\[ \rightarrow \text{ : Forward F.T., } f(x) \rightarrow F(u) \text{ means } \]
\[ F(u) = \int_{-\infty}^{\infty} f(x) \exp(j2\pi ux)dx. \]

\[ \leftarrow \text{ : Inverse F.T., } F(u) \leftarrow f(x) \text{ means } \]
\[ f(x) = \int_{-\infty}^{\infty} F(u) \exp(-j2\pi ux)du. \]
The Fourier transform is an important mathematical tool, applicable to many branches of engineering and science. For example, Fourier transforms are important in the theories of such diverse subjects as signal processing, image processing, antenna theory, x-ray metallography, x-ray crystallography, spectroscopy, optics and radio astronomy. In these fields situations often arise in which it is more convenient to measure the Fourier transform of a quantity rather than the quantity itself. Although both the modulus and phase of the Fourier transform (F.T.) are usually required, the phase may not be obtainable or, if it is, will often not be as easily or as accurately measured as the modulus. When the available phase information is of lower quality than the available modulus information we shall say that a "phase problem" exists. This thesis contributes towards the solution of "phase problems" and two particular ones are examined in detail.

The first problem examined is that of determining the aperture distribution of an antenna from its far-field radiation pattern, without requiring the direct measurement of the phase of the radiation pattern. The techniques described in this thesis which enable this to be done represent a potentially useful contribution to antenna measurement practice because the phase of the radiation pattern of an antenna is nearly always more difficult to measure than the modulus. A planar source distribution can always be obtained from its radiation pattern using the well known F.T. relationship, but in general the phase as well as the modulus
of the pattern is required in the computation of the transform. As was first recognized by Gabor (1948) with his invention of holography, this situation is changed if the source distribution being observed is composed of two separate parts, the unknown source and a known reference source. The concept of applying the principle of holography to the measurement of antenna radiation patterns was suggested by Bates (1971). In this thesis, it is demonstrated, using the results of experiments with acoustic antennas, that the holographic approach leads to pattern measurement techniques that are simple, accurate and practical. It is shown both that Bates' suggestions can be simplified considerably if a point reference source is used and that, in practice, such a reference source can be provided very simply. Several extensions to the basic holographic approach are described and experiments are reported which show them to be practical. Most useful of these extensions are two techniques which allow the holographic approach to be used in the near field.

The second problem examined is the problem of obtaining accurate phase information when radio or optical interferometers are used to obtain maps of the brightness temperature distributions of astronomical sources. The interferometer measures the F.T. of the brightness temperature distribution but in most cases the phase cannot be measured as accurately as the modulus. New, simple, computational procedures are presented which allow the accuracy of the measured phase to be determined. For those cases where the modulus information is more accurate than the phase, a new technique is presented
which provides an improved estimate of the brightness temperature distribution. Detailed structure of the distribution is obtained from the modulus information only, the less accurate phase information being used only to resolve ambiguities. With the current trend in radio astronomy towards shorter wavelengths and longer baselines, accurate phase information will become progressively more difficult to obtain, so the new techniques could find wide application. When no phase information at all is available, for example when a radio or optical intensity interferometer is used, all the possible brightness temperature distributions that could give rise to the measured modulus can be computed using the technique suggested by Bates (1969a). Without some a priori information there is no basis for choosing one possibility rather than another. Bates' technique is used here for the first time to process actual measured astronomical data (taken from the literature) and is shown to be practical. In the special case in which a source that consists of two separated parts is observed, it is shown that there is a possibility of determining the brightness temperature unambiguously from a modulus-only interferometer measurement. The deconvolution technique that is developed in the course of examining this problem is new and its advantages over conventional deconvolution procedures could make it applicable to a wide variety of problems. New material is included in Chapters 2-4 and 6-8, with Chapters 1 and 5 being introductory and Chapter 9 giving conclusions and suggestions for future research.
In order to show the value of techniques that allow the aperture distribution of an antenna to be obtained from the modulus of a far field radiation pattern, Chapter 1 discusses both the value of being able to determine an antenna's aperture distribution and the problems associated with measuring the phase of radiation patterns directly. Previous attempts to reconstruct source distributions from modulus-only pattern measurements are reviewed.

Chapter 2 presents the basic theory for a holographic approach to radiation pattern measurement. The Chapter opens with a brief review of applications of conventional microwave and acoustical holography to the problem of obtaining the field distribution in the aperture of an antenna. The general theory of the new pattern measurement techniques follows the original theory given by Bates (1971). The theory of extensions to the basic technique is described as are the simplifications that result if a point reference source is used. The sources of error in the holographic pattern measurement (H.P.M.) techniques are discussed and a simple practical method of determining the accuracy of a measurement is described.

Chapter 3 presents the results of a series of experiments with acoustic antennas which demonstrate the practical application of the H.P.M. techniques. Detailed, step-by-step procedures are given for the various H.P.M. techniques used. The accuracy of the results obtained demonstrates that H.P.M. techniques are useful for determining the aperture distributions of individual antennas or for determining the modulus, phase and position of the elements of an array of antennas.
Because the H.P.M. techniques are based on the F.T. relationship between the aperture distribution of an antenna and its far-field radiation pattern, they are not directly applicable to near-field pattern measurements. It is shown in Chapter 4 that for the special case of narrow beam antennas whose radiation patterns are effectively confined to a limited range of angles, a F.T. relationship exists between a modified aperture distribution and the near-field radiation pattern. Although this F.T. relationship is the theoretical basis for a defocusing technique devised by Bickmore (1957) to enable the far-field pattern of an antenna to be measured directly in the near-field, it does not appear to have been previously used to compute the far-field pattern from a measured near-field pattern. It is shown, by processing data taken from the literature, that this new F.T. method is a practical method of predicting the far-field from a near-field pattern and has the advantage of being much simpler than other computational procedures. Because the method is based on a F.T. relationship, the H.P.M. techniques can be used to enable the far-field radiation pattern of an antenna to be predicted from a modulus-only measurement in the near-field. A representation for the radiation pattern of an antenna that is applicable in the near field and in the far-field is the angular Fourier series representation. An H.P.M. technique based on this representation is also explained and illustrated in Chapter 4. The technique allows the phase of the radiation pattern of an antenna to be obtained on a circle of any diameter enclosing the antenna.
Chapter 5 introduces the concept of an entire function of exponential type and its zeros in the complex plane. In several measurement situations, notably in antenna radiation pattern measurement and radio astronomical interferometry, the measured quantity can be considered to be the F.T. of a function that is of finite extent (bounded or compact support). In this case, the measured quantity is an entire function of exponential type whose modulus and phase is characterized by the positions of its zeros in the complex plane. The theory of such functions is described and ways in which their zeros can be used in the solution of "phase problems" is discussed. Previous applications of the complex zero theory to "phase problems" and related topics are reviewed and a list of useful theorems (taken from the literature) concerning the zeros is included.

In Chapter 6 the theory of Chapter 5 is applied to obtain solutions to phase problems that occur in astronomical interferometry. The suggestion that this theory be applied to astronomical problems is due to Bates (1969a). The basic theory of astronomical interferometry is discussed and the difficulties of obtaining accurate phase information are outlined. The procedures presented for determining the accuracy of measured phase information and for correcting the phase are new. The results obtained when they are used to process both ideal (computer generated) and measured data (taken from the literature) demonstrate that the techniques are practical and useful. When modulus-only information is available, there are in general many possible brightness temperature distributions that could give rise to the measured interferogram modulus. The results obtained using Bates' (1969a) method for computing all these possible distributions
indicate that, in practice, the physical constraint that the
distributions be wholly positive considerably decreases the
number of possibilities.

Further applications of the complex zero theory are
given in Chapter 7. It is shown that the holographic and
complex zero approaches in combination lead to new and useful
source reconstruction procedures. In particular, the new
deconvolution procedure using complex zeros has several
advantages over conventional deconvolution procedures.

In Chapter 3 the computational techniques used in Chapters
3, 4, 6 and 7 are discussed. All computations were performed
on the IBM 360/44 computer of the University of Canterbury
Computer Centre. With the exception of the function minimiz-
ation routine (section 8.2.1.4) written by Chan (1970) and
the standard IBM subroutines, HARM for the fast Fourier
transform algorithm, POLRT for computing the zeros of poly-
nomials (modified to handle polynomials with complex
coefficients) and PMPY for multiplying polynomials, all
computer programs used were written by the author in Fortran
IV.

Publications to date on topics relevant to this thesis
are:

NAPIER, P.J. (1971), "Array Element Modulus and Phase from
Measured Radiation Pattern Modulus", Proc. IREE Aust.,
accepted for publication.

NAPIER, P.J. and BATES, R.H.T. (1971), "Holographic Approach to
Radiation Pattern Measurement. II. Experimental Verific-


Papers submitted for publication include:


CHAPTER 1. Antenna Aperture Distributions and Radiation Pattern Phase

In Chapters 2, 3 and 4 new measurement techniques are described that allow the aperture distribution of an antenna to be determined from a radiation pattern, without requiring the phase of the pattern. In this Chapter, to establish the value of such techniques, the usefulness of knowing the aperture distribution of an antenna is discussed and the problems associated with measuring the phase of an antenna's radiation pattern are described.

1.1 Identification of Errors in Antennas

An important, and in many cases the most important, design criterion for a directional antenna is that it should have a particular far-field radiation pattern. As a result, the problem of predicting the effect of errors in the structure of an antenna on the radiation pattern of the antenna is one that has received considerable attention in the literature. For single antennas, for example, errors such as the incorrect positioning of a feed source in a large reflecting antenna (Rusch and Potter 1970, p.88), aperture blocking by feed supports (Silver 1949, p.190), reflections from ground planes (Balanis, 1970) and random distortions in a reflector surface (Bates, 1959; Ruze, 1966), have been analysed for their effects on antenna patterns. For antenna arrays, errors such as incorrect modulus and phase of elements (Kummer, 1966) and mutual coupling effects between array elements (Oliner and Malech, 1966) have been analysed to determine their effect on the radiation pattern.
Consider now the inverse problem of relating a particular fault, such as an incorrect beamwidth or sidelobe level, in a measured antenna pattern to the error in the antenna that is causing it. This is a very practical problem but only a few special cases have received attention in the literature. Mott et al. (1965, 1966a) and Cottony (1966) examine the problem of locating reflecting obstacles that are introducing errors into the measured pattern of the antenna and obtain results that have application only in special cases (see section 1.5). Blum et al. (1961) process the radiation pattern of an array to isolate phase errors in the individual elements of the array (see section 3.3.4). The lack of widely applicable techniques for relating particular pattern errors to their causes is not surprising since the general solution of the problem would constitute a solution to the inverse diffraction problem. Consider a closed surface that completely encloses a region of sources whose distribution is unknown. Suppose that the field radiated by these sources is measured over the closed surface. Then the inverse diffraction problem is defined here as the problem of determining the distribution of the sources, given only this measured field pattern. It is only necessary to invoke the equivalence principle (Harrington 1961, p.106) to show that in general there can be no unique solution to this problem. In general a non-denumerable infinity of source distributions can give rise to the same radiation pattern.

Even if the actual source distribution on the structure of an antenna could be determined uniquely from the measured radiation pattern, its usefulness for locating errors in the
antenna structure is doubtful since the required source distribution is unlikely to be known for other than very simple radiating structures. A much simpler and more useful concept than the actual source distribution of an antenna is its equivalent planar source distribution, usually known as the "aperture distribution". This can be related uniquely to the radiation pattern by the well known Fourier transform technique, as is discussed below.

1.2 The Aperture Distribution and Radiation Pattern of an Antenna

1.2.1 The Fourier Transform Relationship

The definition of an aperture distribution of an antenna and its Fourier transform (F.T.) relationship with the far field radiation pattern is now a basic concept for antenna engineers. Woodward and Lawson (1948) and Booker and Clemmow (1950) examine the subject in detail for two dimensional situations and Brown (1958) and Rhodes (1964) treat three dimensional cases. Consider Fig. 1.1 in which an antenna is represented as a volume, V, of monochromatic (radian frequency \( \omega \), wavelength \( \lambda \)) radiating sources existing to the left of the \( x-y \) plane. Rhodes (1964) shows that the field everywhere in the region \( z > 0 \) is determined completely by the values of \( E_x(x,y) \) and \( H_x(x,y) \). \( E_x(x,y) \) and \( H_x(x,y) \) are the \( x \) components of the electric and magnetic fields on the \( xy \) plane. If the \( xy \) plane is situated close to the antenna, for example at the mouth of a horn or at the physical aperture of a reflector antenna, the \( xy \) plane is called the aperture plane of the antenna and the distributions of \( E_x \) and \( H_x \) are known as the aperture distributions of the \( y \) directed equivalent magnetic
and equivalent electric current sources respectively (Jordan 1950, p. 563). The Fourier transforms of these aperture distributions define angular spectrums of plane waves that determine the field in \( z > 0 \). As Booker and Clemmow (1950) point out, it is only when the aperture distribution is effectively limited to a finite part of the complete \( xy \) plane that the concept of a polar diagram or far field radiation pattern is useful. When the aperture distribution is limited in this way the radiation pattern of particular field components at distances large compared to the maximum extent of the aperture distribution is simply the angular spectrum defined by the aperture distribution. Consider, in particular, the case in which the radiation pattern of the antenna in Fig. 1.1 is measured in the \( xz \) plane. Let the aperture distribution of \( y \) directed electric current sources be \( f(x,y) \). Define an equivalent line source distribution \( f(x) \) by

\[
f(x) = \int_{-\infty}^{\infty} f(x,y) \, dy \quad (1.1)
\]

The infinite limits in (1.1) can be replaced by finite limits when the aperture distribution is limited to a finite part of the \( xy \) plane. In the far-field, a suitable expression for \( E_y \) measured at the point defined by the cylindrical coordinates \((R, \theta)\) in Fig. 1.1 is

\[
E_y(R,u) = cF(u) \exp(-jkR)/R \quad (1.2)
\]

where \( c \) is a constant and \( k = 2\pi/\lambda \).

\[
u = \sin \theta / \lambda \quad (1.3)
\]
In (1.2) \( F(u) \) is the far-field radiation pattern of \( E_y \).

Simplification of the expressions given by Rhodes (1964) shows that

\[
F(u) = \int_{-\infty}^{\infty} f(x) \exp(j2\pi ux) dx
\]

(1.4)

\[
f(x) = \int_{-\infty}^{\infty} F(u) \exp(-j2\pi ux) du
\]

(1.5)

which means that \( F(u) \) and \( f(x) \) are a F.T. pair. By measuring \( F(u) \), \( f(x) \) can be obtained straightforwardly using (1.5). The F.T. relationship has been described in terms of the \( y \) directed electric sources in the aperture and the \( y \) component of the electric field in the far-field. F.T. relationships exist between other field quantities as well. Table 1.1 lists the possible aperture quantities that can be obtained by Fourier transforming various radiated field components. In Table 1.1, \( J_y \) and \( J_x \) are the \( y \) and \( x \) components of any actual electric current that exists on the \( xy \) plane, as opposed to \( H_x \) and \( H_y \) which represent the equivalent electric currents (Jordan 1950, p.563). \( J_x \) and \( J_y \) have been included so that

<table>
<thead>
<tr>
<th>( F(u) ) Quantity Measured in Far Field</th>
<th>( f(x) ) Aperture Quantity Obtained by F.T.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_y ) or ( E_\theta )</td>
<td>( E_x )</td>
</tr>
<tr>
<td>( E_y ) or ( H_\theta )</td>
<td>( J_y ) or ( H_x )</td>
</tr>
<tr>
<td>( H_y / \cos \theta ) or ( E_\theta / \cos \theta )</td>
<td>( J_x ) or ( H_y )</td>
</tr>
<tr>
<td>( E_y / \cos \theta ) or ( H_\theta / \cos \theta )</td>
<td>( E_y )</td>
</tr>
</tbody>
</table>

Table 1.1: Far-field and aperture quantities related through the F.T.
Table 1.1 can be used to interpret measurements made on antennas such as linear antennas which do not have a well defined aperture distribution but for which the far-field radiation pattern is nevertheless the F.T. of the source distribution. In Table 1.1 the arrow indicates Fourier transformation as defined in the glossary.

1.2.2 Determination of the Aperture Distribution

In most cases the antenna engineer has to rely heavily on past experience when deciding what corrections to make to an antenna that has an error in its radiation pattern. However, knowledge of the actual aperture distribution of the antenna will often be a useful aid in isolating the fault. For example, errors such as aperture blocking, reflector profile distortion and scattering objects outside the aperture will usually produce errors in the aperture plane distribution that are sufficiently localized to indicate the position of the error, whilst feed displacement is characterized by a phase error across the whole aperture. For an array, a knowledge of the aperture distribution of an array would immediately indicate what changes are required in the excitation of each element to produce the correct modulus and phase distribution.

The aperture distribution of an antenna can be measured directly by scanning a sufficiently small probe across the aperture, as has been demonstrated by Andrews (1950), Ehrlich et al. (1955) and Buchsbaum et al. (1955). However, these references show that such a measurement is very tedious, especially if the whole aperture is to be scanned and if both electric and magnetic field is to be measured in both modulus
and phase. The field probe always distorts the field being measured and the design of a probe that introduces sufficiently small errors is difficult, especially at short wavelengths. Cornbleet (1968) has proposed an interesting alternative to this direct measurement method. He partially obscures the aperture of the antenna by a reflectionless dielectric sheet of known small phase insertion delay. He shows that a precise measurement of the resulting aberration of the antenna's main beam can be used to determine the aperture distribution. The aberrations due to theoretical aperture distributions are calculated and compared with the measured aberration to give the actual aperture distribution so that the method is unlikely to be useful for determining small, unknown errors in the aperture distribution. In many practical situations computing the aperture distribution from the far-field pattern will be preferable to either direct measurement or Cornbleet's (1968) method.

The F.T. relationship between aperture distribution and far-field radiation pattern described in section 1.2.1 has been given only for the case in which the radiation pattern is measured in a single plane. This is the most practical type of pattern measurement. F.T. relationships exist for the complete two dimensional case (Rhodes, 1964) but radiation patterns are seldom measured in sufficient detail over a complete hemisphere to allow the two dimensional F.T. to be carried out. In those cases when the complete two dimensional pattern measurement is made, the techniques of Chapters 2, 3 and 4 can be applied. In the one dimensional case, only the equivalent line source, defined by (1.1), is obtained
rather than the complete two dimensional aperture distribution. Nevertheless, the equivalent line source can still be a valuable aid for isolating errors, especially if it is first found along one axis and then along a perpendicular axis. For linear antennas or linear arrays, the equivalent line source is all that is needed.

The major disadvantage of the F.T. method of obtaining an antenna aperture distribution is that both the modulus and phase of $F(u)$ are required to calculate $f(x)$ using (1.5) (the ambiguity concerning the detailed shape of $f(x)$ when only $|F(u)|$ is known is considered in Chapter 5). In general, the phase of $F(u)$ is much more difficult to measure directly than is the modulus of $F(u)$.

1.3 Measurement of Radiation Pattern Phase

The techniques for measuring the radiation pattern of an antenna are well established (Silver 1949, Chap. 15; Montgomery 1947, Chap. 15; Jasik 1961, Chap. 34; IEEE 1965). The test facility required to measure the pattern of an antenna in both modulus and phase is shown schematically in Fig. 1.2. The distance $R$ at which the pattern should be measured depends on the accuracy required of the measurement (Montgomery 1947, p.574). If $L$ is the width of the test antenna, a distance $R = 2L^2/\lambda$ is usually considered sufficient.

Whilst the measurement of the modulus of the pattern is relatively straightforward and requires little equipment, the need for phase information complicates the measurement considerably. The three pieces of equipment shown in Fig. 1.2 whose provision is most likely to be inconvenient are the
rotating joint, the phase comparator and the R.F. link between opposite ends of the test range. These three items are needed only for the phase measurement. Accurate rotating joints are expensive, especially at very short wavelengths. The actual phase comparison can be made accurately with several different types of bridge circuit (Dyson, 1966) but, again, commercial instruments are expensive and may not be available (Knott, 1969). Providing a link between opposite ends of the range to supply the reference to the phase comparator can be extremely difficult. The electrical length of the link must be stable to a very small fraction of a wavelength over a total distance that can be many thousands of wavelengths. A radiated reference signal must be used if the range is very long (IEEE, 1965).

It is assumed that the radiation pattern phase is required for the purpose of computing the aperture distribution of an antenna. As shown above, the phase measurement is more difficult than the modulus measurement. Thus, any new pattern measurement techniques that allow the aperture distribution to be obtained from only modulus measurements will represent a useful contribution to antenna measurement practice. Such techniques are described in Chapters 2, 3 and 4. Knowledge of the phase of a radiation pattern is also necessary for predicting far-field radiation patterns from near-field measurements (see section 1.4) or for determining the phase centres of antennas (Teichman, 1970). The new techniques of Chapters 2, 3 and 4 are also of value in these cases.
1.4 Radiation Pattern Measurement in Near Field

The problem of predicting the far-field radiation pattern of an antenna from its near-field pattern has received, and continues to receive, considerable attention in the literature. Comprehensive reviews of the subject have been published recently (Jamieson, 1970; Jensen, 1970). Although approximate results can be obtained from only the modulus of the near-field pattern (Bates and Elliot, 1956), accurate methods of computing the far-field from a measured near-field pattern require both the modulus and phase of the near field. Because the distances between the transmit and receive ends of the antenna range are shorter for the near field measurement, the reference signal for the phase comparator can be more easily provided than in the far-field case. The measurement of phase in the near-field will therefore be easier than in the far-field. However, accurate methods of obtaining this phase information from modulus only measurements will still be of value. Techniques described in Chapter 4 contribute towards this for two particular near-field - far-field techniques.

The method of Brown & Jull (1961) is one near field technique that is suitable for obtaining phase information from a modulus only measurement. Brown and Jull represent the field radiated by an antenna by a complete set of orthogonal cylindrical wave functions and complex modal coefficients. By measuring the near-field pattern around a cylindrical surface enclosing the antenna, the measured field is expressed as a Fourier series in these wave functions. The unknown coefficients are then determined by Fourier
analysis of the measured field. Since a complete knowledge of the modal coefficients at a point in the region exterior to the surface which circumscribes the antenna serves to define the fields everywhere in that region, these coefficients specify the total far-field radiation pattern. In section 4.3.3 it is shown that these coefficients can be determined from a modulus only pattern measurement.

The pattern measurement techniques described in Chapters 2 and 3 are based on the F.T. relationship between aperture distribution and far-field-radiation pattern that was discussed in section 1.2.1. It is shown in section 4.3.1 that in certain cases, notably for narrow beam antennas, a similar F.T. relationship holds for distances considerably less than $2L^2/\lambda$. This relationship, which does not appear to have been used previously for computing the far-field from a measured near-field pattern, provides a very simple solution to the near-field - far-field problem. Because of the existence of the F.T. relationship, the techniques of Chapters 2 and 3 can be applied to allow a modulus only measurement to be made in the near-field. This possibility is described more fully in section 4.3.1.

1.5 Previous Users of Modulus Only Pattern Information

In the next three chapters useful experimental techniques are presented that allow the aperture distributions of antennas to be obtained from only the modulus of a radiation pattern. Previous work in this field is limited: Mott et al. (1965) use the modulus of a radiation pattern to determine the importance and position of pattern disturbing structures near an antenna. They show that the F.T. of the modulus of the perturbed pattern gives a source distribution that has sources at the correct positions
of the disturbing elements, but includes spurious sources at incorrect positions. To obtain the result it is necessary to assume that both antenna and disturbing structures are effectively point sources and that the antenna is much stronger than the disturbing sources. Using this method it is not possible to determine whether obstacles are to the left or right of the antenna. In a later letter (Mott et al., 1966a) similar theory is used to show that information about the source distribution can be obtained from the modulus of the pattern whenever the source distribution consists of a dominant central influence and smaller perturbing terms outside the central region. A further letter (Mott et al., 1966b) generalizes the result to two dimensions. In these three reports Mott et al. have made use of the fact that, although little can be said about a single source when the modulus of its pattern is known, the situation is changed if the source consists of two or more separated parts. In this respect the methods are similar to the techniques described in Chapters 2, 3 and 4.

Cottony (1966) also considers the problem of identifying scatterers in radiation pattern measurements. He carries out a simple analysis of the grating lobe structure that is superimposed on the measured radiation pattern by the scatterers to find probable locations of the obstacles. The process is approximate and does not provide a unique answer.

Finally we mention the work of Knott (1969) who obtains the phase of a scattering pattern from modulus only measurements. The technique, which requires three separate patterns to be recorded, two of them having a reference signal added
to the measured signal, is related to the method described in section 4.3.3 as is noted there.
FIGURE 1.1. CO-ORDINATES FOR APERTURE PLANE AND RADIATION PATTERN.

FIGURE 1.2. TEST ARRANGEMENT FOR MEASUREMENT OF MODULUS AND PHASE OF RADIATION PATTERN.

This chapter presents the basic theory of new antenna radiation pattern measurement techniques which allow the aperture distribution of an antenna to be obtained from only the modulus of a measured radiation pattern. The techniques are based on the principle of holography. Conventional applications of microwave and acoustic holography to antenna measurements are reviewed to indicate in what way the new techniques differ from conventional holography.

2.1 Introduction

It was suggested in Chapter 1 that the aperture distribution of an antenna can be a useful aid in isolating errors in an antenna. The phase of the radiation pattern of the antenna that is usually needed to obtain the aperture distribution is difficult, or at least inconvenient, to measure directly. In this chapter the basic theory of antenna pattern measurement techniques that allow the aperture distribution to be obtained from only the modulus of a pattern is given. These techniques are called holographic pattern measurement (H.P.M.) techniques.

For the purpose of the present work "holography" is defined as "any measurement procedure that allows complete modulus and phase information about the field radiated by a coherent source to be obtained from the measurement of only the modulus of a field". The concept of holography was originally introduced by Gabor (1948) in the field of electron microscopy. More recently it has been very widely applied in optical, acoustical and microwave situations. The very large amount of literature on holography currently
appearing in the journals (Kallard, 1970) and the availability of several modern texts (Stroke, 1966; DeVelis and Reynolds, 1967; Goodman, 1968) make it unnecessary to repeat basic holographic theory here. Conventional optical holography can conveniently be classified into four types: Fresnel holography, Fraunhofer holography, Fresnel or Fraunhofer sideband (off axis) holography and Fourier transform holography; the theory of, and differences between, these various types are given by DeVelis and Reynolds (Chapter 2, 1967). As explained below in section 2.3.4, the H.P.M. techniques can be considered to be particular examples of Fourier transform holography.

2.2 Non-Optical Holography

Some of the H.P.M. techniques described later in this Chapter are related to holographic methods already used in microwave and acoustic holography. Previous work in these fields is now reviewed.

2.2.1 Microwave Holography

The first reported use of holography using microwave radiation is the work of Dooley (1965) who used Gabor's (1948) original form of Fresnel holography to obtain an image of a metal shape. Later workers have used Fresnel off-axis (Aoki, 1969) and Fraunhofer (Swingler, 1970) holograms to obtain images of reflecting objects. In these measurements a hologram suitable for optical reconstruction is formed by modulating the intensity of a light source by a signal proportional to the intensity of the microwave field measured by a probe moved over the hologram plane. The position of the light source is slaved to the position of the probe and
the hologram is recorded photographically. Methods of
recording a microwave hologram directly without the need for
a scanned light source are being developed by using liquid
crystals as the recording medium (Gregoris and Iizuka, 1970).
The usual method of reconstructing an image from the micro-
wave holograms is to treat suitably scaled versions of the
holograms as normal optical holograms and to reconstruct them
optically. Reconstruction by a computer operating on a
sampled version of the hologram has been carried out (Aoki
and Boivin, 1970), although the improvement over optical re-
construction is small.

The examples of microwave holography described above
are, like most optical applications, cases in which the object
being imaged is "passive" and can be imaged only in so far as
it reflects or diffracts the radiation from a coherent source.
Holography can also be used to obtain an image of an active
source of coherent radiation, offering a means of determining
the source distribution of an antenna.

Deschamps (1967) notes that holography could be used "to
observe the currents on a transmitting antenna" and points
out several important differences between optical and micro-
wave holography. The most important difference is that at
microwave frequencies the field incident on the hologram plane
can be coherently detected whilst at optical frequencies only
intensity detectors are available. This means that the
reference wave need not be actually radiated but can be added
electronically to the signal measured at the hologram plane.
By suitably phase shifting the reference signal, an "off axis
hologram" can be obtained allowing separation of the
reconstructed images and the zero order terms. Rope and Tricoles (1969) and Swingler and Anderson (1970) using this phase shifted reference signal technique, and Farhat and Guard (1970) using a Fourier transform hologram technique have obtained microwave holograms and optical reconstructions of the field distributions in antenna apertures. Although the phase of the aperture distributions is contained in such optical reconstructions, there is no way of measuring it, and the modulus information is generally rather poor. In an excellent mathematical treatment of holographic imaging Mittra and Ransom (1967) discuss the reasons for the generally poor quality of images obtained using microwave holography.

Because objects imaged by microwaves are generally much smaller in terms of wavelengths than objects imaged optically, they radiate over much wider angles so that more information is lost by measuring the hologram over only a small part of the hologram plane. Further loss of resolution is inevitable when an equivalent optical transparency of the hologram is constructed from the microwave hologram because of the very large scaling factor required to account for the difference in optical and microwave wavelengths.

Iizuka and Gregoris (1970) propose microwave holography as a useful aid in the analysis and design of antennas. Their holographic measurement of a radiating monopole emphasizes an important difference between conventional microwave holography and the H.P.M. techniques described below. Iizuka and Gregoris reconstruct the field radiated by a monopole and show that the field appears to come from point sources situated only at the ends of the monopole, even though
the actual source distribution is known to be a continuous sinusoidal distribution (Jordan 1950, p.314). Because the H.P.M. techniques are based on the F.T. relationships between far-field and source distribution, Table 1.1 shows that measurement of the correct far-field quantity would allow the actual currents on the monopole to be obtained. The point is that conventional microwave holography only reconstructs fields whilst the H.P.M. techniques allow actual source distributions to be obtained, provided the sources are related to the far-field pattern through a F.T.

Finally we mention the interesting possibility of reconstructing microwave holograms using microwave radiation. Checcacci et al. (1968, 1971) and Tricoles and Rope (1967) have constructed microwave holograms of antennas using microwave absorbent materials which, when reconstructed using microwave radiation, reproduce the radiation behaviour of the original antennas with surprisingly good accuracy.

2.2.2 Acoustic Holography

The application of holographic imaging principles to acoustical imaging closely parallels the development of microwave holography. Acoustic wavelengths commonly used for holography are similar to those used in microwave holography so that experimental arrangements are almost the same. As with microwave holography any of the lensless forms of holography may be used acoustically (Wetherell et al., 1967). The advantage of being able to coherently detect the acoustic signals, as in the microwave case, has led to the use of electronically injected reference waves (Massey, 1963) and other methods of solving the multiple image problem (Skattebol,
Photographic transparencies of the acoustic holograms are usually made using the scanned light source technique although direct methods of formation have been investigated (Berger H. 1969). Source distributions for acoustic radiators have been obtained by Greene (1969) but the accuracy of his results would be insufficient to detect anything other than very large errors in a source distribution. Most applications of acoustic holography are concerned with the imaging of reflecting objects.

As with microwave holography, most acoustic holograms are reconstructed optically although computer reconstruction is possible when sufficient computer storage is available (Goodman, 1969). The problems with acoustic holography are the same as those discussed for microwave holography in section 2.2.1 with the additional problem that the acoustic propagation medium (usually air or water) is less stable than in the microwave situation.

2.3 A Holographic Approach to Antenna Radiation Pattern Measurement

2.3.1 Basic Theory

It has been shown in section 1.2.1 that the effect of any source distribution can be represented exactly by radiations from an equivalent planar source distribution known in radio engineering as the "aperture distribution". To determine the aperture distribution of an antenna from its radiation pattern, both the modulus and phase of the pattern must be measured. However, this is not the case when the radiating source distribution is a mixture of known (reference) and unknown sources. A hologram may be regarded as a record
of the modulus of the radiation pattern of a mixture of known and unknown sources and from such a record the unknown source distribution may be reconstructed. The application of this holographic principle to the measurement of antenna radiation patterns and aperture distributions was originally proposed by Bates (1971). His work provides the basis for the rest of this chapter.

Using the definitions of section 1.2.1, denote by \( f(x) \) and \( F(u) \) respectively the equivalent line source distribution and far-field radiation pattern of a radiating source. \( f(x) \) and \( F(u) \) are a F.T. pair (equations (1.4), (1.5)). Define \( R(u) \) to be the intensity of the radiation pattern.

\[
R(u) = F(u)F^*(u) = |F(u)|^2
\]  

(2.1)

Bates (1971) defines a "Fraunhofer hologram" to be a record of \( R(u) \) when \( f(x) \) is a mixture of a known reference source \( g(x) \) and an unknown source \( h(x) \). Thus

\[
f(x) = g(x) + h(x)
\]  

(2.2)

so that, assuming a linear medium and by an obvious extension of notation

\[
F(u) = G(u) + H(u).
\]  

(2.3)

Substituting (2.3) into (2.1) gives

\[
R(u) = G(u)G^*(u) + H(u)H^*(u) + G(u)H^*(u) + H(u)G^*(u)
\]  

(2.4)

Consider only source distributions that are of finite extent (that is \( f(x) \) is a function with compact, or finitê, support).
Define the positions and extents of \( f(x) \), \( g(x) \), \( h(x) \) by:

\[
f(x) = 0 \text{ if } x \text{ is outside the range } A < x < A + L \tag{2.5}
\]

\[
g(x) = 0 \text{ if } x \text{ is outside the range } A_1 < x < A_1 + L_1 \tag{2.6}
\]

\[
h(x) = 0 \text{ if } x \text{ is outside the range } A_2 < x < A_2 + L_2 \tag{2.7}
\]

Denote by \( L_0 \) the greater of \( L_1 \) and \( L_2 \). Thus

\[
L_0 = L_1, \quad L_1 > L_2 \tag{2.8}
\]

\[
L_0 = L_2, \quad L_1 < L_2 \tag{2.9}
\]

When \( g(x) \) and \( h(x) \) do not overlap, denote their separation by \( D \). Using (2.5) and (2.7)

\[
D = A_2 - A_1 - L_1, \text{ } g(x) \text{ to left of } h(x) \tag{2.10}
\]

\[
D = A_1 - A_2 - L_2, \text{ } g(x) \text{ to right of } h(x) \tag{2.11}
\]

As an example of the notation, Fig. 2.1 illustrates definitions (2.5)-(2.11) for the particular parameter values given in Table 2.1 in units of wavelengths, when \( g(x) \) and \( h(x) \) are uniform distributions and \( g(x) \) is to the right of \( h(x) \).

<table>
<thead>
<tr>
<th></th>
<th>( f(x) )</th>
<th>( h(x) )</th>
<th>( g(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A )</td>
<td>( L )</td>
<td>( A_2 )</td>
<td>( L_2 )</td>
</tr>
<tr>
<td>-4</td>
<td>10.5</td>
<td>-4</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2.1: Parameter values for example of Fig. 2.1.
Consider now a measurement of $R(u)$ made with a known $g(x)$ and with the separation $D \gg L_0$. Such a measurement could be made, for example, with the experimental arrangement shown in Fig. 2.2 which shows a plan view of two antennas mounted on a turntable with their apertures in the same plane passing through the centre of the turntable. The distance $R$ is greater than the far-field distance of the source distribution ($R \gg 2L^2/\lambda$). Notice that it may be more convenient to measure $|F(u)|$ and to obtain $R(u)$ from (2.1) rather than to measure $R(u)$ directly. Assuming that $f(x)$ is as shown in Fig. 2.1, then $|F(u)|$ is shown in Fig. 2.3.

Define $r(x)$ as the F.T. of $R(u)$

$$r(x) = \int_{-\infty}^{\infty} R(u) \exp(j2\pi ux)du \quad (2.12)$$

Then, using the convolution theorem for Fourier transforms (Bracewell 1965, Pg. 112) (2.4) gives

$$r(x) = f(x) \otimes f^*(x)$$

$$= g(x) \otimes g^*(x) + h(x) \otimes h^*(x) + g(x) \otimes h^*(x) + h(x) \otimes g^*(x) \quad (2.13)$$

where the symbol $\otimes$ denotes correlation and is defined in the glossary. Fig. 2.4 shows the $r(x)$ corresponding to $f(x)$ of Fig. 2.1 and $|F(u)|$ of Fig. 2.3. Since $D \gg L_0$, (2.6) to (2.13) show that

$$r(x) = f(x) \otimes f^*(x) = 0$$

if $x$ is in the range $L_0 < |x| < D \quad (2.14)$

so that, as demonstrated in Fig. 2.4, $r(x)$ separates on inspection into three distinct parts. The part centred at
x = 0 is the autocorrelation term \[ g(x)g^*(x) + h(x)h^*(x) \]. Because of the separation D between \( g(x) \) and \( h(x) \), the cross correlation terms in (2.13) are displaced from the origin by the distance D. Notice that the three parts of \( r(x) \) only separate completely if D is greater than either of the extents of \( g(x) \) and \( h(x) \). Define \( e(x) \) as the outside part of \( r(x) \) to the right.

\[ e(x) = r(x), \text{ for } x \text{ in the range } x > D \] (2.15)

Comparing (2.15) with (2.13) and using the definition of correlation contained in the glossary shows that

\[ e(x) = g(x)g^*(x), \text{ } g(x) \text{ to left of } h(x) \] (2.16)

\[ e(x) = h(x)h^*(x), \text{ } g(x) \text{ to right of } h(x) \] (2.17)

Define \( E(u) \) to be the inverse F.T. of \( e(x) \). Applying the convolution theorem for Fourier transforms to (2.16) and (2.17) gives

\[ e(x) \rightarrow E(u) = G(u)H^*(u), \text{ } g(x) \text{ to left of } h(x) \] (2.18)

\[ e(x) \rightarrow E(u) = H(u)G^*(u), \text{ } g(x) \text{ to right of } h(x) \] (2.19)

where the arrows denote Fourier transformation as defined in the glossary. Since \( g(x) \) is known, \( G(u) \) can be computed using \( g(x) \rightarrow G(u) \). Then, from (2.18) and (2.19) the unknown radiation pattern, \( H(u) \), may be obtained as

\[ H^*(u) = E(u)/G(u), \text{ } g(x) \text{ to left of } h(x) \] (2.20)

\[ H(u) = E(u)/G^*(u), \text{ } g(x) \text{ to right of } h(x) \] (2.21)

The unknown aperture distribution, \( h(x) \), can now be computed
using \( H(u) \rightarrow h(x) \). If, as will usually be the case, it is known whether \( g(x) \) is to the left or right of \( h(x) \), then (2.20) or (2.21) can be used directly to obtain \( H(u) \) and \( h(x) \). If the relative positions of \( g(x) \) and \( h(x) \) are not known, \( h(x) \) can still be found as follows. Let \( L_3 \) be the width of \( e(x) \). Then (2.6), (2.7), (2.16) and (2.17) show that

\[
L_3 = L_1 + L_2
\]  

(2.22)

Since \( g(x) \) is known, \( L_1 \) is known and \( L_3 \) can be obtained by examining \( e(x) \). Then (2.22) gives \( L_2 \). Define \( E_1(u) \) and \( E_2(u) \) as

\[
E_1(u) = E(u)/G(u), \quad E_2(u) = E(u)/G^*(u)
\]  

(2.23)

and \( e_1(x) \) and \( e_2(x) \) as the inverse F.T. of \( E_1(u) \) and \( E_2(u) \) respectively

\[
E_1(u) \rightarrow e_1(x), \quad E_2(u) \rightarrow e_2(x)
\]  

(2.24)

Usually \( e_1(x) \) and \( e_2(x) \) will have different widths. In this case \( h(x) \) can be chosen as the member of the pair \( e_1(x), e_2(x) \) that has width \( L_2 \). If \( e_1(x) \) and \( e_2(x) \) have the same width then Bates (1971) shows that \( e_1(x) \) and \( e_2(x) \) give \( h(x) \) and \( h^*(x) \) although it is not possible to tell which is which.

In summary, in this section it has been shown that an unknown source distribution can be obtained from only the modulus of a radiation pattern. This radiation pattern must be the pattern of the unknown source combined with a known reference source and the two sources must be separated by a distance greater than either of their extents. This technique will be called the holographic pattern measurement
(H.P.M.) technique with arbitrary reference source and complete separation.

2.3.2 H.P.M. Techniques with Partial Separation

An essential part of the method described above is that the condition \( D > L_0 \) must be satisfied. Only when the known and unknown sources are separated by a distance greater than either of their extents does \( r(x) \) separate by inspection into three parts, allowing the crosscorrelation term \( g(x)@h^*(x) \) or \( h(x)@g^*(x) \), necessary for reconstructing \( h(x) \), to be obtained. We now consider the case in which the separation is not greater than either of the extents of \( g(x) \) and \( h(x) \), that is \( 0 < D < L_0 \). Methods used in this case will be called H.P.M. techniques with arbitrary reference source and partial separation.

Consider first the case in which the width of the known reference source, \( g(x) \), is greater than the width of the unknown source, \( h(x) \) (note that this is not a very practical situation for reasons described in Section 3.2.2). Assume that the separation between the two sources is greater than \( L_2 \) but less than \( L_1 \)

\[
L_1 > L_2 \quad (2.25)
\]

\[
L_1 > D > L_2 \quad (2.26)
\]

In this case (2.6), (2.7) and (2.13) show that if \( e(x) \) is now defined as

\[
e(x) = r(x) - g(x)@g^*(x), \quad x > D \quad (2.27)
\]

then (2.16) and (2.17) hold and \( h(x) \) can be obtained by operating on \( e(x) \) using (2.18)-(2.21). Since \( g(x) \) is known,
$g(x)@g^*(x)$ can be easily computed. As an example of this H.P.M. technique with partial separation, suppose that $h(x)$ in Fig. 2.1 is shifted to have its centre at $x = -1$ so that (2.25) and (2.26) hold. Fig. 2.5 shows $r(x)$ and $r(x) - g(x)@g^*(x)$ for this case. Fig. 2.5 shows that $h(x)@g^*(x)$ can be obtained by inspection from $r(x) - g(x)@g^*(x)$.

When the separation between $h(x)$ and $g(x)$ is less than either of their extents, but greater than zero, the method given immediately above can be used in cases in which the reference signal is very much larger than the unknown one so that $h(x)@h^*(x)$ is negligible compared to $h(x)@g^*(x)$. In this case, (2.27) is approximately satisfied allowing $h(x)$ to be computed with an accuracy that improves as the relative strength of $g(x)$ with respect to $h(x)$ increases. If $h(x)@h^*(x)$ is not negligible, $e(x)$ must be defined as

$$e(x) = r(x) - g(x)@g^*(x) - h(x)@h^*(x), \quad x > D \quad (2.28)$$

(2.28) requires that $|H(u)|$, the modulus of the radiation pattern of the unknown source, be known so that $h(x)@h^*(x)$ can be computed using $|H(u)|^2 \rightarrow h(x)@h^*(x)$. $|H(u)|$ can be readily obtained by switching off the reference antenna. Then $h(x)$ can be found by processing $e(x)$ defined in (2.28) in the same way as in section 2.3.1.

When $D < 0$, and $h(x)$ and $g(x)$ actually overlap, $h(x)$ cannot be obtained easily from $R(u)$ unless $h(x) \ll g(x)$. In this case the method presented in section 7.2.2 may be used. Another possibility is a method requiring three patterns to be measured, with the phase of the reference being altered between measurements. Such a method for the special case in
which \( g(x) \) is a point source is described at the end of section 2.3.3.

### 2.3.3 H.P.M. Techniques with Point Source Reference

All the H.P.M. techniques described in sections 2.3.1 and 2.3.2 can be greatly simplified if the reference source, \( g(x) \), is a point source. Assume that

\[
g(x) = S\delta(x+T_R)
\]  

(2.29)

where \( \delta(x) \) is the Dirac delta function, \( S \) is the strength of the point source situated at \( x = -T_R \). If (2.29) is satisfied then (2.13), which can be regarded as the basic equation for the H.P.M. technique, becomes

\[
r(x) = S^2\delta(x) + h(x)\delta h^*(x) + S h(x+T_R) + S^* h(-T_R-x)
\]  

(2.30)

Notice that the last two terms on the R.H.S. of (2.30) give the unknown source, \( h(x) \), directly. This means that the H.P.M. techniques described in the two previous sections all give \( h(x) \) without requiring the processing to remove \( h(x) \) from \( h(x)\delta g^*(x) \) or \( g(x)\delta h^*(x) \). Further, those methods that required subtraction of \( g(x)\delta g^*(x) \) are simplified since the autocorrelation of \( g(x) \) is merely a delta function that can be neglected.

In the case when the point source and \( h(x) \) overlap, three separate pattern measurements are required. These are: 
- \( |H(u)| \), the pattern of the test antenna by itself; 
- \( |F_1(u)| \), the pattern of the test antenna and point source together; 
- \( |F_2(u)| \), the pattern of the test antenna and point source together with the phase of the point source advanced by \( \pi/2 \).
The three quantities $h(x)\otimes h^*(x)$, $r_1(x)$ and $r_2(x)$ can now be computed where

$$|H(u)|^2 \rightarrow h(x)\otimes h^*(x) \quad (2.31)$$

$$|F_1(u)|^2 \rightarrow r_1(x) = S^2\delta(x) + h(x)\otimes h^*(x)$$
$$+ Sh^*(x-T_R) + S^*h(-T_R-x) \quad (2.32)$$

$$|F_2(u)|^2 \rightarrow r_2(x) = S^2\delta(x) + h(x)\otimes h^*(x)$$
$$+ jSh^*(x-T_R) - jS^*h(-T_R-x) \quad (2.33)$$

Then, neglecting $S^2\delta(x)$, the autocorrelation of the reference source, which can be removed by inspection, the relationship

$$[r_1(x) - h(x)\otimes h^*(x)] - j[r_2(x) - h(x)\otimes h^*(x)] = 2Sh^*(x-T_R) \quad (2.34)$$

gives the unknown source distribution.

2.3.4 Interpretation of H.P.M. Techniques

It will be useful to point out to what extent the radiation pattern measuring techniques, discussed theoretically in sections 2.3.1-2.3.3 and applied experimentally in Chapter 3, can be considered "holographic". The techniques certainly satisfy the broad definition of holography given in section 2.1. In similarity with conventional holographic techniques the measured hologram is the intensity of the combined radiations of a known and an unknown source. In conventional holography a Fourier Transform hologram is one in which the fields recorded by the hologram are the Fourier transforms of the unknown and reference source distributions (Develis and Reynolds 1967, Chap. 2). The H.P.M. techniques can be placed in this category. The particular property of a Fourier Transform hologram of both reconstructed images being
real images is evident in the H.P.M. techniques since both of
the cross correlation terms $h(x)@g^*(x)$ and $g(x)@h^*(x)$ lead to
$h(x)$.

The single most important difference between the H.P.M.
techniques and the examples of holography given in sections
2.2.1 and 2.2.2 is that the H.P.M. techniques are one-diimen-
sional whereas conventional microwave and acoustic holography
is two-dimensional. Rather than being a limitation, it is
this fact that makes the H.P.M. techniques practical and
widely applicable. As will be seen in Chapter 3, application
of H.P.M. techniques presents no more difficulties than are
normally encountered when measuring the modulus of the
radiation pattern of an antenna. Practical application of
conventional microwave holography is restricted to small
antennas. It will be observed that the reconstructed
aperture distributions obtained in Chapter 3 are considerably
better in general than the optically reconstructed distribut-
ions reported in the references of sections 2.2.1 and 2.2.2.
Furthermore, conventional optical reconstruction does not
give the phase of the aperture distribution. With the H.P.M.
techniques, because the hologram is reconstructed in the
computer, both modulus and phase information is available.

2.4 Accuracy of H.P.M. Techniques

2.4.1 A Criterion for Estimating the Accuracy of the H.P.M.
Techniques

An essential part of any useful measurement technique is
a means of estimating the error in the measurement. A valid
criterion which may be used to guage the accuracy of a
holographic pattern measurement is now discussed.

Fourier transforms of measurable functions are unique so that there is only one \( f(x) \), having finite measure (functions of zero measure are uninteresting physically), which could produce a given \( F(u) \). But if only \(|F(u)|\) is given, it is well known (see section 5.1.3) that there are in general many \( f(x) \) which could produce it. However, in the H.P.M. techniques investigated here, the a priori information consists of more than just \(|F(u)|\). Both the modulus and phase of \( G(u) \), the radiation pattern of the reference source, are known as well. In this case, both the modulus and phase of \( H(u) \) are heavily constrained by the forms of the last three terms on the right hand side of equation (2.4). This suggests a procedure for determining the accuracy of any particular reconstruction of the aperture distribution of an antenna under test.

(a) Measure \(|F(u)|\).

(b) Measure \(|H(u)|\). This may be simply accomplished by removing the feed to the reference antenna.

(c) Reconstruct \( h(x) \) in the manner required for the particular H.P.M. being used.

(d) Compute \( H(u) \) by Fourier transformation of \( h(x) \) using \( h(x) \rightarrow H(u) \).

(e) Observe the difference \( \epsilon(u) \) between the two \(|H(u)|\) obtained in (b) and (d)

\[
\epsilon(u) = |H(u)|_{\text{measured}} - |H(u)|_{\text{reconstructed}} \tag{2.35}
\]

If \( \epsilon(u) \) is small, then the method may be considered accurate. The point is that there are many functions \( H(u) \) which could satisfy the right hand side of equation (2.13) given only
|F(u)| and G(u). But there is only one such function, given G(u) and |H(u)|. Thus, the modulus and phase of H(u) are unique so that, by the uniqueness of Fourier transforms, h(x) is unique.

There are two main sources of error when the H.P.M. techniques are used to obtain an antenna's aperture distribution. These are pattern truncation and measurement noise.  

2.4.2 The Effect of Pattern Truncation

It is not physically possible to measure the radiation patterns of sources in the interval |sinθ| > 1.

If \( R_m(u) \) represents the intensity pattern actually measured, a suitable expression for \( R_m(u) \) is

\[
R_m(u) = R(u) \text{ rect}(u\lambda/2) + N_R(u)
\]

(2.36)

where \( \text{rect}(u) \) is the rectangular gate function defined in the glossary and \( N_R(u) \) is measurement noise. Since the accuracy of the reconstructed source distribution is dependent on the accuracy with which \( r(x) \) can be obtained, it is necessary to examine the effect that this truncation has on \( r(x) \). Consideration of the effects of measurement noise is deferred until section 2.4.3.

The problem of computing the F.T. of truncated data arises in many measurement techniques. Source reconstruction by interferometry (Bracewell and Roberts, 1954), spectral analysis by Fourier spectroscopy (Steel 1967, p. 191), calculation of frequency responses from time domain measurements (Bates and Burroll, 1970) and examination of crystal structure by x-ray diffraction (Ramachandran and Srinivasan, 1970) are all affected by the necessarily finite length of the measured
data. The reason for the truncation may be a limitation of the measuring apparatus, such as an interferometer which will only be able to take measurements for baselines less than some maximum value or, as in the present case, the data may be physically unmeasurable for a certain range of the measurement variable. The radiation pattern cannot be measured in the far field for the range of (complex) angles, sometimes known as the "invisible angles", for which 

\[ |\sin(\theta)| > 1 (|u| > 1/\lambda) \]

The magnitude of the radiation pattern in the "invisible region" determines the reactive power in the aperture (Rhodes, 1964). Rhodes shows that for apertures of finite extent the radiation pattern is analytic for all finite values of \( u \) and the pattern in the invisible region can in principle be obtained as the analytic continuation of the part of the pattern measured in visible space.

The possibility of extending the range of truncated data by analytic continuation of this sort has received attention in such fields as optical image restoration (Andrews 1970, p.24, considers this problem and reviews the relevant literature), spectroscopy (Williams and Chang, 1966) and radio astronomy (Schell, 1965 and 1969; Biraud, 1969). The amount of extrapolation that can be obtained using such techniques is often negligible in practice and is always limited by the amount of noise in the available data. Data extrapolation was not attempted in the present work.

A qualitative measure of the effect on \( r(x) \) of truncating \( R(u) \) can be obtained from (2.36). Let \( r_m(x) \) be the estimate of \( r(x) \) obtained by Fourier transforming \( R_m(u) \). Then, using the convolution theorem for the F.T.
where the symbol \( \ast \) denoting convolution and \( \text{sinc}(x) \) are defined in the glossary. The effect of the convolution with \( \text{sinc}(2\pi x/\lambda) \) is to smooth out features in \( r(x) \) and to increase its extent. Since \( \text{sinc}(2\pi x/\lambda) \) has a main lobe of width \( \lambda/2 \), the resolution in \( r_m(x) \) is no better than \( \lambda/2 \). The other effect of the convolution with \( \text{sinc}(2\pi x/\lambda) \) is the well known "Gibbs phenomenon" which is the introduction of spurious oscillations into \( r_m(x) \) at points where \( r(x) \) changes rapidly. This effect can be reduced by appropriate data weighting during the computation of \( r_m(x) \), as described in section 8.1.

It is very difficult to place quantitative bounds on the errors in \( r_m(x) \) due to the truncation in \( R_m(u) \) since the errors depend directly on those portions of the radiation pattern that have not been measured. Bates (1971) has attempted to place an upper bound on the truncation error. In the present notation, Bates' (1971) expression gives

\[
|r(x) - r_m(x)|_{\text{max}} \approx \frac{\overline{R}_m(-1/\lambda) + \overline{R}_m(1/\lambda)}{\pi x} \tag{2.38}
\]

where \( \overline{R}_m(u') \) is the peak value of the lobe of \( R_m(u) \) situated closest to \( u = u' \). \( |r(x) - r_m(x)|_{\text{max}} \) represents the maximum possible truncation error in the estimate of \( r(x) \). The R.H.S. of (2.38) increases rapidly as \( x \) gets smaller. It has been found that (2.38) sometimes gives an excessively large estimate for the error. For the measurements reported in the following Chapters, it was assumed that truncation effects were included in the composite noise level proposed below in section 2.4.4.
2.4.3 The Effect of Measurement Noise

An additional source of error in \( r_m(x) \), the computed estimate of \( r(x) \), is the unavoidable presence of measurement noise represented by \( N_R(u) \) in (2.36).

If the statistics of \( N_R(u) \) are known an estimate of the error in \( r_m(x) \) due to this measurement noise can be made. For example, if \( N_R(u) \) can be assumed to be zero-mean additive Gaussian white noise, an error estimate derived by Smith (1969) can be used. In the present notation, Smith shows that if \( N \) samples of \( R_m(u) \) are taken at intervals of \( \Delta u \), then the error in \( r_m(x) \) due to the noise will exceed a value \( \varepsilon \) for a fraction \( \xi \) of the time where

\[
\varepsilon = \Delta u \sqrt{2N\Phi(0)} \operatorname{erf}^{-1} \left[ \sqrt{1-\xi} \right] \quad (2.39)
\]

In (2.39) \( \Phi(x) \) is the autocorrelation of \( N(u) \) and \( \operatorname{erf}^{-1}(x) \) is defined in the glossary. For example, using this expression Nicholson (1968) shows that with \( \xi = 0.01 \), the noise amplitude \( \varepsilon \) in \( r(x) \) exceeded at not more than 1% of the points calculated in the \( x \) domain is

\[
\varepsilon = 1.99\Delta u \sqrt{2Nd^2} \quad (2.40)
\]

where \( d^2 \) is the variance of the samples of \( R_m(u) \) due to measurement noise. Expression (2.40) gives a useful estimate of the noise error for regions of the \( x \) domain in which the noise is constant (white noise) and \( d^2 \) is a measure of the short-term variance of the noise in the \( u \) domain.

The error estimate in (2.40) may be applied directly if \( N_R(u) \) in (2.36) is Gaussian white noise and if an estimate of the variance of \( N_R(u) \) can be obtained. When, as in the
measurements reported in Chapters 3, 4 and 7, the modulus, 
$$|F_m(u)|$$, rather than the intensity, $$R_m(u)$$, of the radiation 
pattern is recorded, the measurement can be described by

$$|F_m(u)| = |F(u)| + N_F(u) \quad (2.41)$$

where $$N_F(u)$$ is the noise in the measured modulus and pattern 
truncation has been neglected for the moment. The intensity of the 
hologram is obtained as the square of this measured modulus

$$R_m(u) = |F_m(u)|^2 = |F(u)|^2 + 2|F(u)|N_F(u) + N_F^2(u) \quad (2.42)$$

In estimating the effect of $$N_F(u)$$ on the estimate of 
r(x) obtained from $$R_m(u)$$, the second order small quantity 
$$N_F^2(u)$$ in (2.42) can be neglected. Notice that (2.42) indicates 
that one of the penalties of measuring only the modulus of 
$$F(u)$$ is a 3db reduction in signal to noise ratio. With $$N_F^2(u)$$ 
assumed negligible, and noting that $$|F(u)| = F(u) \exp(-\arg F(u))$$ 
the F.T. of $$R_m(u)$$ in (2.42) gives

$$r_m(x) = r(x) + 2f(x)*p(x)*N_F(x) \quad (2.43)$$

where

$$\exp(-\arg F(u)) \longrightarrow p(x) \quad (2.44)$$

and

$$N_F(u) \longrightarrow n_F(x) \quad (2.45)$$

The second term on the R.H.S. of (2.43) represents the error in $$r_m(x)$$. 

In most cases $$N_F(u)$$ may be considered to be white 
Gaussian noise in which case (2.45) may be used to place an 
upper limit on $$n_F(x)$$. It is difficult to place a limit on
\( f(x) \ast p(x) \ast n_F(x) \). However, as shown in Bates (1971), \( p(x) \) will usually be of infinite extent and the effect of the double convolution in \( f(x) \ast p(x) \ast n_F(x) \) is to spread and smooth \( n_F(x) \) throughout the whole \( x \) domain. This fact provides the basis for the choice of a practical noise level in the computed estimate of \( r(x) \), as is described below.

### 2.4.4 A Practical Noise Level for \( r(x) \)

The H.P.M. theories described in sections 2.3.1-2.3.3 all require that the source distributions involved in the measurements be of finite extent. In practice, actual antenna aperture distributions are not identically zero outside the physical aperture of the antenna, but have fringing fields of extent comparable to at least a wavelength that makes their boundary fuzzy. However, for the types of antennas for which the H.P.M. techniques will be most useful, that is antennas with large apertures and narrow main beams, these fringing fields will be very small compared to the fields within the aperture, so that the assumption of finite aperture size is a very practical one.

All the H.P.M. techniques rely on the fact that \( r(x) \), or \( r(x) \) with parts subtracted out, separates into three parts which can be identified by inspection. Even though the source distributions can be considered to be of finite extent, the discussion in sections 2.4.2 and 2.4.3 shows that in any practical situation \( r_m(x) \) will not separate exactly into three parts because it is contaminated by errors due to pattern truncation and measurement noise. In practice, a threshold level must be applied to \( r_m(x) \) to separate it into three parts. All parts of \( r_m(x) \) having magnitude less than
the threshold level are set to zero. Denote by $V$ the smallest threshold level that must be applied to $r_m(x)$ to separate it into three parts which are of finite extent. Bates (1971) suggests that an estimate of $V$ can be obtained from a knowledge of the noise level in $R_m(u)$ and the behaviour of $R_m(u)$ in the region $u = u_{\text{max}}$. However, as was mentioned in section 2.4.2, Bates’ estimate of the truncation error sometimes gives an estimate of $V$ that is excessively high. A simpler method of determining $V$ that has been found adequate in practice is to choose $V$ by examining $r_m(x)$. When, as in most measurements, the source widths $L_1$ and $L_2$ and the separation $D$ between the sources are known, $V$ can be chosen as the minimum threshold necessary to set $r_m(x)$ to zero in the ranges $L_0 < |x| < D$ and $|x| > D + L_1 + L_2$. For example, when the ideal $r(x)$ shown in Fig. 2.1 is contaminated by measurement noise and truncation error it could typically appear as shown in Fig. 2.6. Since it is known (Table 2.1) that $L_2 = 2$, $L_1 = 4$ and $D = 4.5$, $r_m(x)$ should be zero for $4 < x < 4.5$ and $x > 10.5$ and $V$ is chosen so as to satisfy these conditions as closely as possible. In Fig. 2.6 the threshold $V$, which is indicated, is determined by the maximum modulus of $r_m(x)$ in the range $x > 10.5$.

In cases where the exact dimensions of the sources or their separation are not known, choice of the best value for $V$ will be more difficult. This problem arises, for example, when unknown reflecting objects near $h(x)$ are expected to introduce significant perturbations into $H(u)$, in which case $L_2$ will not be known exactly. Suppose that the $r_m(x)$ shown in Fig. 2.7 were to be computed from a measured $R_m(u)$. In
this case it would not be clear whether $V_1$ or $V_2$ shown in Fig. 2.7 is the correct threshold. Consequently, in cases where the choice of $V$ is ambiguous, the data should be processed using several values of $V$ to determine which one leads to the best agreement between the reconstructed $|H(u)|$ (the F.T. of the reconstructed $h(x)$) and the $|H(u)|$ actually measured.

The value chosen for $V$ gives an estimate of the error level in $r_M(x)$. If the object of making an H.P.R. measurement is to locate errors in $h(x)$, perturbations in $h(x)$ that are smaller than $V$ cannot be considered to be significant. As explained in section 2.4.1, the most important check on the accuracy of the reconstructed $h(x)$ is the agreement obtained between the reconstructed and measured values of $|H(u)|$. 
FIGURE 2.1. EXAMPLE OF APERTURE DISTRIBUTIONS \( g(x), h(x) \).

FIGURE 2.2. EXAMPLE OF MEASUREMENT OF \( R(u) \).

FIGURE 2.3. \( |F(u)| \) FOR \( F(x) \) SHOWN IN FIGURE 2.1.
FIGURE 2.4. \( r(x) \) FOR \( f(x) \) SHOWN IN FIGURE 2.1.

FIGURE 2.5. --- \( r(x) \) AND -------\( [r(x) - g(x) \odot g^*(x)] \) FOR THE \( f(x) \) SHOWN IN FIG. 2.1, WITH \( h(x) \) CENTRED AT \( x = -1 \).

FIGURE 2.6. TYPICAL EXPERIMENTAL ESTIMATE FOR \( r(x) \) OF FIGURE 2.1.
FIGURE 2.7: \( r_m(x) \) IN WHICH CHOICE OF CORRECT THRESHOLD \( V \) IS AMBIGUOUS.
CHAPTER 3. A Holographic Approach to Radiation Pattern Measurement – Experiment

The basic theory of a holographic method for measuring antenna radiation patterns and aperture distributions has been given in the previous chapter. In this chapter, experimental applications of the theory are described.

3.1 Experimental Arrangement

In order to test the theory of the H.P.M. techniques it is necessary to use sources which are related to their radiation patterns by Fourier transformation. The far-field radiations from electromagnetic or acoustic antennas satisfy this requirement. Because of their ready availability in our laboratory, acoustic antennas were used in the experiments reported here. The experimental arrangement used for the measurements described in this chapter is shown schematically in Fig. 3.1. Brief details about the pieces of equipment shown in Fig. 3.1 are given in Appendix 1. The measurements were made in air using continuous wave transmission at an acoustic wavelength, \( \lambda \), of 1 cm, (the frequency was approximately 34.5 kHz). Solid dielectric electrostatic transducers (Martin, 1969) made in the Electrical Engineering Department of the University of Canterbury were used as sources. The particular transducers used in each experiment are described in the appropriate sections below. Radiation patterns were recorded on a penrecorder fed by a linear detector and a potentiometer mounted on the turntable shaft.
For the measurements described in Section 3.2, the reference antenna was mounted on the turntable with the test antenna. In this case, the reference antenna and the test antenna were positioned on the turntable with their vibrating surfaces arranged to be in a plane containing the axis of rotation of the turntable. They were offset from the axis by distances $T_R$ and $T_T$ respectively. Notice that for the measurements of section 3.2, the fixed reference antenna was not included in the arrangement.

For the measurements of section 3.3, the reference signal was provided by the fixed reference antenna that remained stationary as the turntable rotated. For these measurements, there was no reference antenna on the turntable and the test antenna was positioned by itself on the turntable in the manner described in the preceding paragraph. The fixed reference antenna was located directly below the turntable and it was strongly decoupled from the test antenna.

The receiving antenna was in the far-field of the combination of test and reference antennas; a separation greater than $2L^2/\lambda$ was used, where $L$ is the effective total extent of the radiating sources. From section 2.3.1, $L = L_1 + L_2 + D$. The experiments were carried out in the middle of a laboratory measuring 60 ft x 40 ft x 15 ft, with a padded floor. The orientation of the antenna range was chosen to minimise wall reflections which did not significantly affect the accuracy of the measurements reported here. The frequency was adjusted between measurements to keep the wavelength constant with temperature (Beranek 1949, p.47). Small thermal variations and drift of the acoustic medium introduced
noise into the measured patterns. To minimise the effects of the noise the average of several recordings was taken for each measurement. The uncertainty introduced into the measured radiation patterns by this noise was typically ±1.5% of the maximum value of the pattern. This uncertainty was estimated by examining the differences between successive recordings of the same pattern.

In all the measurements described, values were computed for both moduli and phases of all correlation functions, aperture distributions and radiation patterns, but for convenience only the moduli are usually shown here (the argument of section 2.4.1 demonstrates the sufficiency of this). In a few cases the phase of $H(u)$ was measured with the phase meter shown in Fig. 3.1 so that the reconstructed phase could be compared with it. The phase meter was used only for these checking measurements.

Radiation patterns are plotted versus $u = (\sin\theta)/\lambda \text{ cm}^{-1}$, where $\theta$ is the angle measured from the perpendicular to the aperture planes of the antennas. Since $\lambda = 1 \text{ cm}$ for the radiation used in the experiments, $u = ±1$ corresponds to $\theta = ±90^\circ$.

3.2 Application of Basic H.P.M. Technique With Complete Separation and Arbitrary Reference Source

3.2.1 Procedure and Results of Experiment

If the aperture distribution, $g(x)$, of an antenna is known then the theory of section 2.3.1 shows that this known antenna can be used as a reference source to determine the aperture distribution, $h(x)$, of an unknown antenna. Consider
the following procedure.

Procedure A - H.P.M. with complete separation and arbitrary reference source.

(A.1) Measure \(|F(u)|\), the modulus of the far-field radiation pattern of \(g(x)\) and \(h(x)\) together, with the separation between \(g(x)\) and \(h(x)\) greater than either of their widths.

(A.2) Compute \(r(x)\) using \(|F(u)|^2 \rightarrow r(x)\).

(A.3) Determine the noise threshold \(V\) (Section 2.4.4) and apply it to \(r(x)\) which then separates by inspection into three parts.

(A.4) Define \(e(x)\) as the right-hand part of \(r(x)\). Compute \(E(u)\) using \(e(x) \rightarrow E(u)\).

(A.5) Compute \(H(u)\) using (2.10) or (2.21). Compute \(h(x)\) using \(H(u) \rightarrow h(x)\).

The results of a practical application of Procedure A are now described. The test and reference antennas used in the experiment were circular transducers with diameters of 1.7 cm and 3.9 cm respectively. \(g(x)\), the aperture distribution of the reference antenna was known, having been obtained using the H.P.M. technique described in section 3.4.2. \(h(x)\), the aperture distribution of the 3.9 cm transducer, was unknown.

The two transducers were positioned on the turntable as described in section 3.1 with the test antenna on the right of the reference antenna (i.e. \(g(x)\) to left of \(h(x)\)). The separation \(D\) between the antennas was 4.5 cm which was greater than either of the extents of \(g(x)\) and \(h(x)\). The modulus, \(|F(u)|\), of the combined radiation pattern of the two antennas is shown in Fig. 3.2. The known pattern of the
reference antenna, $G(u)$, is also shown in Fig. 3.2. The phase of $G(u)$ was linear to within $\pm 4^\circ$. Fig. 3.3 shows $r(x)$, the Fourier transform of $|F(u)|^2$. As predicted in section 2.3.1, because of the separation between the antennas, $r(x)$ divides into three separate parts. The small non-zero values between, and outside, these parts were treated as noise (c.f. section 2.4.4) and all values with modulus below the threshold $V$ shown in Fig. 3.3, were set to zero. (2.16) shows that $e(x)$, the part of $r(x)$ for which $x > 4$, is the cross correlation $g(x) \otimes h^*(x)$. $|H(u)| = |E(u)/G(u)|$ is shown in Fig. 3.4 together with the actual measured $|H(u)|$ for comparison. The reconstruction is seen to be almost perfect over most of the interval $|u| < 1$. Fig. 3.5 shows $h(x)$, obtained from $H(u)$ through $H(u) \rightarrow h(x)$

3.2.2 Discussion of Basic H.P.M. Technique

The crux of the reconstruction procedure used above is the deconvolution step required to obtain $h(x)$ from $g(x) \otimes h^*(x)$ when $g(x)$ is known. The successful computation of $H(u)$ from $E(u)/G(u)$ is dependent on $G(u)$ being carefully chosen. If $G(u)$ is zero for any values of $u$ then the division cannot be carried out unless $E(u)$ is exactly zero for the same values of $u$. In general, $E(u)$ will never be exactly zero because of errors in $e(x)$ (remember that $e(x)$ has a noise level of approximately $V$). When $G(u)$ is not exactly zero but is very small the division process will lead to large errors in $H(u)$ unless $G(u)$ is accurately known. The problem of de-correlation (or deconvolution) is considered further in sections 7.4.1-7.4.3 in which relevant literature is reviewed and some solutions to the problem proposed, but it is useful at this point to illustrate the difficulties that arise if
g(x) is not carefully chosen.

In a measurement prior to the one described in section 3.2.1 a 1 cm diameter circular transducer was used as the reference antenna with the test antenna the same as in section 3.2.1. e(x) was similar to that shown in Fig. 3.3 except that it was narrower and the noise threshold \( V \) was twice as large. In this measurement \( G(u) \) was measured directly by switching off \( h(x) \) and exciting only \( g(x) \). \(|G(u)|\) is shown in Fig. 3.6. Significant perturbations, due primarily to reflections from the antenna mounts and the test antenna can be seen on the edges of \(|G(u)|\). The result of computing \(|H(u)| = |E(u)/G(u)|\) in this case is also shown in Fig. 3.6 and the effects of the errors in \( G(u) \) are clearly seen. The \( h(x) \) reconstructed from this \( E(u) \) is shown in Fig. 3.7. \( h(x) \) contains large errors. The improvement obtained in the measurement described in section 3.2.1 results mainly from the better choice of \( G(u) \). The 1.7 cm antenna radiated less in the direction \( u = 1 \) so that reflections from the test antenna were reduced. Notice, however, that because \( E(u) \) must be divided by \( G(u) \), \( g(x) \) cannot be so wide that \( G(u) \) has sidelobes at angles for which \(|H(u)|\) still has sensible value.

Another advantage of choosing \( g(x) \) correctly is that often only \(|G(u)|\) will need to be known since, for a well designed antenna, the phase of \( G(u) \) will be constant over most of the main beam.

In summary then, the reference antenna should be chosen smaller than the test antenna to ensure that \( G(u) \) does not have nulls for values of \( u \) for which \( H(u) \) is significant, but it should not be so small that \( G(u) \) is significant in the
direction of the test antenna.

When measuring the radiation patterns of large antennas it is usually not possible to avoid the effects of reflections from objects which intercept the radiation emitted or received by the antennas. If, as often happens, the unwanted reflections are localized, then the theory of section 2.3.1 and Procedure A can be applied: \( g(x) \) and \( h(x) \) would represent the radiating sources in respectively a standard antenna and the pattern disturbing structures. This method would appear to be more powerful than the methods of Mott et al. (1965, 1966a, 1966b) and Cottony (1966) that were discussed in section 1.5.

Because of the inherent difficulty in the deconvolution step, it is not expected that the H.P.M. technique with arbitrary reference source will be the most widely applicable H.P.M. technique for obtaining the aperture distribution and modulus and phase of the radiation pattern of an antenna. Procedure A does have the advantage that no rotating feed joint is required during the measurement of \( |F(u)| \), since the transmitter or receiver can be placed on the turntable with the two antennas. However, the greater simplicity and accuracy of the H.P.M. methods described in sections 3.4.2-3.4.3 will usually make them preferable. A final disadvantage of Procedure A is that \( |F(u)| \) must be measured in the far-field of \( f(x) \), that is, in the far-field of \( g(x) \) and \( h(x) \) together. Since the width of \( f(x) \) must be at least twice the width of \( h(x) \) (see Fig. 2.1), an antenna range will be required that is at least four times as long as would be needed to measure \( h(x) \) by itself.
3.3 Application of H.P.M. Techniques to Measurement of Array Aperture Distributions

The H.P.M. techniques that use arbitrary reference sources can be usefully applied to the problem of determining the modulus and phase of the elements in an array. These applications are described below in sections 3.3.1-3.3.4.

3.3.1 The Calibration of an Array

The successful operation of an antenna array is highly dependent on the accuracy of the modulus and phase of each element. Errors in element modulus and phase produce unwanted sidelobes, reduction in gain and collimation errors (Morimoto and McLean, 1967; Kummer, 1966). Techniques for calibrating the modulus and phase of array elements exist (Swarap and Yang, 1961; Morimoto and Labrum, 1967; Little, 1969). After the array has been calibrated, the radiation pattern may still be found to contain errors. In this case it will not be clear how to relate these pattern errors to errors in the gain, phase or position of the elements of the array and to determine what corrections are needed. With the methods described in the next two sections, the gain, phase and position of the array elements can be found directly from the radiation pattern of the array, allowing differences between design and actual values to be detected rapidly. The techniques are similar to a method successfully used by Blum (1961), as is explained in section 3.3.4.

In principle, once the far-field radiation pattern of the array has been measured in both modulus and phase, its F.T. gives the aperture distribution of the array. If it is made at all, the phase measurement is likely to be less
accurate than the modulus measurements, for the reasons given in section 1.3. However, since an array is composed of many separate antennas, by appealing to the theory of the H.P.M. techniques it should be possible for one of these antennas to act as a phase reference for the rest, thereby allowing all required information to be obtained from only the modulus of the radiation pattern.

3.3.2 Application of H.P.M. With Complete Separation

Consider a linear array of \( N \) identical radiating elements, each element having an aperture distribution of the same shape, call it \( c(x) \). If the aperture distribution of the whole array is \( a(x) \), then

\[
a(x) = \sum_{n=1}^{N} b_n c(x+d_n)
\]  

(3.1)

where \( b_n \) is a complex number indicating the gain and phase of the \( n^{th} \) element and \( d_n \) is the position of the \( n^{th} \) element. Let \( d_1 = 0 \). Choose the element at the right of the array to be element 1 and order the remaining elements according to distance from element 1. There is no requirement that the elements be uniformly spaced. Fig. 3.8 shows a schematic diagram of an array and its aperture distribution.

Consider element 1 to be the reference antenna for a holographic pattern measurement. In H.P.M. nomenclature

\[
g(x) = b_1 c(x)
\]  

(3.2)

Now divide the remainder of the array into two halves and switch off that half that is adjacent to \( g(x) \). Call the other half of the array \( h(x) \). Thus, with \( M \) the smallest integer for which \( d_M > \frac{1}{2} d_N \), set

\[
b_n = 0, \quad M > n > 2
\]  

(3.3)
and \( h(x) \) is given by
\[
h(x) = \sum_{n=M}^{N} b_n c(x+d_n) \tag{3.4}
\]
g(x) and \( h(x) \) are now separated by a distance greater than either of their extents, so that the H.P.M. technique with arbitrary reference and complete separation can be applied directly to obtain \( h(x) \). If \(|F(u)|\), the modulus of the radiation pattern of \( g(x) \) and \( h(x) \) combined is measured, then as described in section 3.2.1 \( h(x) \otimes g^*(x) \) can be obtained from the F.T. of \(|F(u)|^2\). From (3.2) and (3.4)
\[
h(x) \otimes g^*(x) = b_1^* \sum_{n=M}^{N} b_n c(x+d_n) \otimes c^*(x) \tag{3.5}
\]
(3.5) shows that \( h(x) \otimes g^*(x) \) is composed of autocorrelations of the element aperture distributions, \( c(x) \otimes c^*(x) \), centred at the points \( x = d_n \) for \( N > n > M \) with the relative amplitudes of the autocorrelations at these points being given by the complex \( b_n \). The gap between the elements of the array is given by \(|d_n - d_{n-1} - L|\), where \( L \) is the width of \( c(x) \). If this gap is at least as large as the width of \( c(x) \), as will often be the case, then these autocorrelation functions will be completely separated and the \( b_n \) and \( d_n \) can be obtained simply by inspection of \( h(x) \otimes g^*(x) \) without the need for any deconvolution or for accurate knowledge of \( c(x) \) and \( b_1 \). This will often be possible when the gap between elements is less than the extent of \( c(x) \) because of the highly peaked nature of \( c(x) \otimes c^*(x) \). If the separation between elements is not sufficient for the individual elements to be separated in \( h(x) \otimes g^*(x) \), \( h(x) \) must be obtained from \( h(x) \otimes g^*(x) \) using the deconvolution step (A.5).
To summarize, the modulus phase and position of half of the elements of a linear array can be obtained using the following procedure.

Procedure B - Modulus, Phase and Position of Half the Elements of an Array.

(B.1) Call the element at the right hand end of an array \( g(x) \). (The changes in the procedure needed to use the left hand element as reference are obvious – see section 2.3.1.)

(B.2) Divide the remainder of the array into two halves and switch off that half adjacent to \( g(x) \). Call the left half of the array \( h(x) \). Denote by \( D \) the separation between \( g(x) \) and \( h(x) \).

(B.3) Measure \( |F(u)| \) the modulus of the radiation pattern of \( g(x) \) and \( h(x) \) combined.

(B.4) Compute \( r(x) \) using \( |F(u)|^2 \rightarrow r(x) \). Determine the noise threshold \( V \) (section 2.4.4) and apply it to \( r(x) \) to obtain \( e(x) \) by inspection where \( e(x) = r(x), x > D \).

(B.5) \( e(x) \) is the crosscorrelation term \( h(x)@g^*(x) \). Examine \( e(x) \) to determine the modulus, phase and position of the elements in \( h(x) \), as explained in the discussion following equation (3.5).

The results of a practical application of Procedure B are now described. The array measured was composed of seven circular transducers, each of 1 cm diameter, arranged in a straight line. The centre to centre spacing of adjacent transducers was 2.5 cm as shown in Fig. 3.9. The transducers were fed in parallel from the same signal source and were mounted so as to have their vibrating surfaces in the same
plane to within ±0.5 mm, giving a possible phase variation between transducers of less than ±18° (electrical phase shifts due to variations in feeder length were negligible at 34 kHz). The relative excitations of the array elements were obtained by measuring the modulus of the radiation pattern of each transducer with the other six transducers unexcited. The relative excitations, normalized with respect to the average excitation, are shown in Fig. 3.9. The relatively large uncertainties in the excitations are caused by a high noise level in the measured patterns of the individual transducers, the noise being due to reflections from adjacent transducers in the array.

To test Procedure B the excitation to elements 2, 3 and 4 was removed and the modulus of the radiation pattern, call it $|F_A(u)|$, of the combination of elements 1, 5, 6 and 7 was measured. $|F_A(u)|$ is shown in Fig. 3.10a. As expected, the F.T. of $|F_A(u)|^2$ divided into three parts. The right hand part $h(x)\Theta e^x(x)$, is shown in Fig. 3.10b. The estimate of the noise level, $V$, is included in Fig. 3.10b. The position of transducers 5, 6 and 7 at 10 cm, 12.5 cm and 15 cm is plainly indicated as are their relative moduli and phases which agree with expected values within the limits of experimental error. The excitation of transducer 6, situated at 12.5 cm, was then reduced to 0.6 of its previous value and its phase was retarded by 52°. Elements 1, 5 and 7 were unchanged. The radiation pattern, $|F_B(u)|$, of the four elements was again measured and was of similar form to $|F_A(u)|$, but with higher sidelobes between the main grating lobes. The outer part of the F.T. of $|F_B(u)|^2$ is shown in Fig. 3.10b, the change in element
6 at 12.5 cm being accurately indicated.

3.3.3 Application of H.P.M. With Partial Separation

Although the method described in Section 3.3.2 is simple and accurate it has the disadvantage that only half of the array can be measured at once. This may be inconvenient if it is expected that mutual coupling between opposite halves of the array is causing errors in the radiation pattern. A method is now described that allows all elements to be measured at once.

As in section 3.3.2, equation (3.2), let \( g(x) \) be element 1 of the array. Let \( h(x) \) be all the rest of the array elements.

\[
h(x) = \sum_{n=2}^{N} b_n c(x+d_n)
\]

(3.6)

In this case, because \( g(x) \) and \( h(x) \) are not separated by a distance greater than the extent of \( h(x) \), when \( r(x) \) is calculated \( h(x)\oplus h^*(x) \) and \( h(x)\oplus g^*(x) \) will overlap so that \( h(x)\oplus g^*(x) \) cannot be obtained directly from \( f(x)\oplus f^*(x) \) as is possible using the H.P.M. technique with complete separation. However, \( [r(x) - h(x)\oplus h^*(x)] \) separates into three parts, as described in section 2.3.2. The centre part is \( g(x)\oplus g^*(x) \) and the outer parts are \( g(x)\oplus h^*(x) \) and \( h(x)\oplus g^*(x) \) to the left and right respectively. Consider the centre and right hand parts

\[
g(x)\oplus g^*(x) + h(x)\oplus g^*(x) = b^*_1 \sum_{n=1}^{N} b_n c(x+d_n)\oplus g^*(x)
\]

(3.7)

As explained for equation (3.5), equation (3.7) shows that examination of the centre and right-hand parts of

\( [r(x) - h(x)\oplus h^*(x)] \) will give the relative positions, moduli
and phases of all elements in the array. This method then requires two radiation patterns to be measured, one of \(|F(u)|\), the whole array, and one of \(|H(u)|\), the whole array without the end element. The F.T. of \(|F(u)|^2 - |H(u)|^2\) then gives the required information. Notice that \(|F(u)|\) and \(|H(u)|\) will be similar, especially for large arrays. Since the required information is carried only in the differences between \(|F(u)|\) and \(|H(u)|\) the patterns must be measured with sufficient accuracy, so that these differences are well above the error level. Because of the subtraction, this technique will be less accurate than the method of section 3.3.2. Notice that for small arrays it may be possible to make \(g(x)\) much stronger than \(h(x)\), in which case \(h(x)\overline{h^*(x)}\) would be negligible and the subtraction unnecessary.

In summary, the following procedure gives the positions, moduli and phases of all the elements of an array.

**Procedure C - Modulus, Phase and Position of All the Elements of an Array**

*(C.1)* Call the element at the right hand end of an array \(g(x)\).

Call the rest of the array \(h(x)\).

*(C.2)* Measure \(|F(u)|\), the modulus of the radiation pattern of the complete array.

*(C.3)* Measure \(|H(u)|\), the modulus of the radiation pattern of the complete array without the right hand element.

*(C.4)* Compute \(r(x) - h(x)\overline{h^*(x)}\) using

\[
\left[ |F(u)|^2 - |H(u)|^2 \right] \longrightarrow \left[ r(x) - h(x)\overline{h^*(x)} \right].
\]

*(C.5)* Examine \(r(x) - h(x)\overline{h^*(x)}\) to obtain the modulus, phase and position of all the elements of the array, as explained above.
For the 7 element array described in section 3.3.2 the pattern \( |F(u)| \) of the whole array is shown in Fig. 3.10a. \( |H(u)| \), the pattern of elements 2, 3, 4, 5, 6 and 7 was similar, with the same grating lobe and slightly different sidelobe structure. Fig. 3.10c shows the F.T. of \( |F(u)|^2 - |H(u)|^2 \). The position of all elements is accurately indicated. The relative excitations of the elements, normalized with respect to the average excitation, are shown in Table 3.1.

<table>
<thead>
<tr>
<th>Element</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excitation</td>
<td>0.96±.1</td>
<td>1.21±.1</td>
<td>1.13±.1</td>
<td>1.15±.1</td>
<td>1.03±.1</td>
<td>0.82±.1</td>
<td>1.01±.1</td>
</tr>
</tbody>
</table>

Table 3.1: Calculated excitations of array elements

These excitations agree with the expected values shown in Fig. 3.9 within the limits of experimental error. The phases of the elements lie within the expected limits of ±180°, although phase errors are evident for the first two elements since each element should have constant phase across its aperture. The errors result from the errors in the measured radiation pattern being magnified by the subtraction step (0.4).

3.3.4 Discussion of the Array Measurement Techniques

The crux of the array measurement methods described above, as with all the H.P.M. techniques, is that \( h(x)g^*(x) \) can be picked out of the F.T. of \( |F(u)|^2 \). If, instead of being added, \( H(u) \) and \( G(u) \) are used as inputs to a multiplier followed by a low pass filter, then \( \text{Re}(H(u)G^*(u)) \) is obtained (Deschamps, 1967), where \( \text{Re} \) denotes the real part.
Since
\[
\text{Re}(H(u)G^*(u)) = \frac{1}{2}(H(u)G^*(u) + G(u)H^*(u))
\] (3.8)
it can be seen that \( h(x)@g^*(x) \) can be obtained by Fourier transforming \( \text{Re}(H(u)G^*(u)) \), provided \( h(x) \) and \( g(x) \) do not overlap. This is the basis for an array calibration procedure suggested and used by Blum et al. (1961). Their method may be used to measure the whole array at once without requiring the pattern subtraction needed in the method of section 3.3.3. It should be remembered that the purpose of calibrating an array is to correct the smallest detectable errors. This means that the simplest and most accurate measurement procedures should be used. The method of section 3.3.2 is certainly simpler than the method of Blum et al., because only the linear detection of a signal, rather than the multiplication of two signals, is required. The methods described in this paper will be most useful when the array pattern is measured with a coherent source with a good signal to noise ratio which is usually the case in conventional antenna pattern measurements. It would also be true for very large radar or radio astronomical arrays measured with an airborne or satellite-borne transmitter or receiver. In such situations, as well as giving the relative modulus and phase of the array elements, Procedures B and C give the positions of the elements and will indicate periodic errors in the array and errors in the individual antenna apertures. The method of Blum et al. cannot give the last three of these pieces of information. If a radio astronomical source is used to
measure the pattern of the array then the multiplicative (or phase switching) receiver has known advantages (Jennison 1966, Chap. 8) and the method of Blum et al. may be preferable.

In the methods described in sections 3.3.2 and 3.3.3, it has been assumed, and demonstrated for the particular experimental arrangement used, that the modulus and phase of the individual components of $h(x)$ can be determined by inspection of $h(x)g^*(x)$, without needing to carry out a decorrelation. If the separation between the array elements is not large enough, or if all array elements do not have the same aperture distribution $c(x)$, $g(x)$ must be known accurately and the decorrelation step used in section 3.2.1 must be applied. The method proposed in section 2.4.1 can be used to determine the accuracy of reconstructions obtained in this way.

The theory as given above is one-dimensional and is therefore applicable directly to determining the modulus and phase of the elements of a linear array when its radiation pattern is measured in a plane containing the axis of the array. It can also be directly applied to two-dimensional arrays. Consider the simple two dimensional circular array shown in Fig. 3.11. The plane in which the radiation pattern is measured intersects the plane of the array at a line passing through the array. Suppose that, in Fig. 3.11, the elements below this line are switched off. Then the equivalent one dimensional aperture distributions of the remaining elements are obtained by projecting the aperture distributions onto this line. If all the individual distributions are separated, as they are in the example of Fig. 3.11, the
techniques can be used to give the modulus and phase of each element. If all the distributions do not separate, several measurements would have to be made.

3.4 H.P.M. Techniques With a Point Reference Source

The chief disadvantage with the H.P.M. technique described in section 3.2.1 was the need to carry out a decorrelation procedure to obtain the unknown aperture distribution. As explained in section 2.3.3, use of a point source as a reference source considerably simplifies the H.P.M. methods because the decorrelation step is not required. In fact, use of a point reference source leads to some extremely simple and accurate H.P.M. techniques.

3.4.1 Provision of a Point Reference Source Using a Fixed Reference Antenna

The crux of the techniques described theoretically in section 2.3.3 is the provision of the point reference source $S\delta(x+T_R)$. The provision of an actual point source situated at $x = -T_R$ on the turntable of Fig. 3.1 would be very difficult in practice. Even if a reference antenna could be found that was sufficiently small to be regarded as a point source over the range $|u| < 1/\lambda$, when it was mounted at $x = -T_R$ on the turntable direct reflections from the test antenna would introduce large errors of the type illustrated in section 3.2.2. However, if the point source is situated at the centre of the turntable, that is

$$g(x) = S\delta(x)$$ (3.9)

a simple solution to the problem is available. A source $S\delta(x)$ will produce a signal that remains constant in modulus.
and phase at the receiving antenna as the turntable rotates. Such a signal can be provided by a fixed antenna, not mounted on the turntable, that remains stationary as the turntable rotates. Thus, the effect of the fixed reference antenna shown in Fig. 3.1 is identical to actually placing an ideal point source at the centre of the turntable. If the test antenna is then placed (by itself) on the turntable with its centre at $T_T$ as shown in Fig. 3.1, and if $T_T > \frac{3L_2}{2}$, $h(x)$ will be separated from $g(x)$ by a distance greater than either of their extents and the H.P.M. technique with complete separation can be applied. If $\frac{3L_2}{2} > T_T > \frac{L_2}{2}$, then partial separation techniques must be used.

3.4.2 H.P.M. Technique With Point Source Reference and Complete Separation

The following procedure may be used to obtain the aperture distribution of an antenna:

Procedure D - H.P.M. Technique With Point Source Reference and Complete Separation

(D.1) Mount the test antenna to be measured on a turntable with the edge of the antenna nearest the axis of the turntable separated from the axis by a distance greater than the width of the antenna.

(D.2) Arrange a fixed reference antenna as shown in Fig. 3.1. The reference antenna can be placed in any position not on the turntable, so long as it is strongly decoupled from the test antenna. Adjust the input to the reference antenna so that it gives a signal at the receiving antenna that is approximately
(D.3) Measure $|F(u)|$, the radiation pattern of the test and reference antennas together.

(D.4) Compute $r(x)$ using $|F(u)|^2 \rightarrow r(x)$. Determine the noise threshold $V$ (section 2.4.4) and apply it to $r(x)$ which then separates into three parts. Define $e(x)$ as the outside right hand part.

(D.5) $e(x)$ gives the unknown aperture distribution $h(x)$ directly since $e(x) = h^*(x)$ if $h(x)$ to right of turntable axis, $e(x) = h(-x)$ if $h(x)$ to left of turntable axis.

Results of two practical applications of Procedure D are now discussed. In the first experiment, a circular transducer with diameter 3.9 cm was placed on the turntable as shown in Fig. 3.1, with $T_T = 6.3$ cm. There were no other antennas on the turntable. $|F(u)|$, the modulus of the pattern of the combination of the test antenna radiation pattern and the constant reference signal from the fixed antenna is shown in Fig. 3.12. Fig. 3.13 shows $r(x)$, the F.T. of $|F(u)|^2$. As expected, the two autocorrelation terms $h(x)@h^*(x)$ and $g(x)@g^*(x)$ (a delta function) are centred at $x = 0$, and the cross-correlation terms, in this case the required aperture distributions themselves, are out to the sides. As described in section 2.4.4 a threshold $V$ had to be applied to separate the three parts. Any features in $h(x)$ which have a modulus smaller than $V$ may not be regarded as being meaningful. The reconstructed radiation pattern, $H(u)$, which is the F.T. of the aperture distribution, is shown in Fig. 3.14. Close agreement with the actual measured pattern is observed,
indicating that the computed distribution is accurate.

In the second measurement, the aperture distribution of the 3.9 cm diameter circular transducer was perturbed by attaching two circular 1 cm diameter transducers to it in the positions shown in Fig. 3.15. The total width of the source was 4.2 cm. The vibrating surfaces of the 1 cm transducers were approximately 1.5 mm in front of the vibrating surface of the 3.9 cm transducer. All three transducers were fed in phase from the same source and h(x) is the equivalent aperture distribution of the combination of the three transducers. \( T_T \) was equal to 7.75 cm. \(|F(u)|\) for this measurement was similar in form to Fig. 3.12 but differed in detail. \( r(x) = f(x)f^+(x) \) is shown in Fig. 3.15. The autocorrelation of the point source has been removed from Fig. 3.15. The required aperture distribution h(x) is that part of r(x) in Fig. 3.15 that is centred on \( x = 7.75 \) cm and is above the threshold \( V \). h(x) clearly indicates the presence of the 1 cm transducers. The phase shift caused by the 1.5 mm displacement between the vibrating surfaces of the transducers is accurately shown. The reconstructed radiation pattern, \( \mathbb{H}(u) \), which is the F.T. of h(x) is shown in Fig. 3.16. The close agreement with actual measured values of \(|\mathbb{H}(u)|\) shows that h(x) is accurate. The technique illustrated above provides a very simple method for determining an aperture distribution from modulus measurements only. In common with the method of section 3.2, however, it has a serious disadvantage for large antennas. Because of the separation required between the test antenna and the reference source, the effective extent of the source being measured is at least twice that of the test
antenna itself. The distance to the far field of \( f(x) \) is thus at least four times that of \( h(x) \) by itself. For large antennas this may be very inconvenient. In this case the method used in the next section may be preferable.

3.4.3 H.P.M. Technique With Point Reference Source and Partial Separation

The following H.P.M. method can be used to obtain the aperture distribution of an antenna from a radiation pattern measured at the far field distance of the antenna.

Procedure E - H.P.M. Technique With Point Source Reference and Partial Separation

(E.1) Mount the test antenna to be measured on a turntable with one edge of the antenna over the centre of the turntable.

(E.2) Set up a fixed reference antenna as described in step (D.2).

(E.3) Measure \( |F(u)| \); the radiation pattern of the test and reference antennas together.

(E.4) Measure \( |H(u)| \); the radiation pattern of the test antenna by itself.

(E.5) Compute \( r(x) \) using \( |F(u)|^2 \rightarrow r(x) \) and compute \( h(x)\#h^*(x) \) using \( |H(u)|^2 \rightarrow h(x)\#h^*(x) \).

(E.6) Compute \( [r(x) - h(x)\#h^*(x)] \). Determine the noise threshold \( V \) (section 2.4.4) and apply it to \( [r(x) - h(x)\#h^*(x)] \) which then separates into three parts. Define \( e(x) \) as the outside right hand part.

(E.7) Obtain the unknown distribution, \( h(x) \), as explained in step (D.5).

Since the effective position of the point reference source is
on the edge of \( h(x) \), the total extent of \( f(x) \) is only \( L_2 \), so \(|F(u)|\) can be measured at a range of \( 2L_2^2/\lambda \).

Results of a practical application of Procedure E are now described. The antenna measured using this technique was the same 3.9 cm diameter circular transducer used in the first measurement of section 3.4.2. The transducer was placed on the turntable with one edge over the centre of the turntable \( (T_T = 1.95 \text{ cm}) \). \(|F(u)|\), the modulus of the pattern of the antenna and the fixed reference source together, is shown in Fig. 3.17. \(|H(u)|\), the modulus of the pattern of the test antenna by itself fed with the same signal strength is also shown in Fig. 3.17. The Fourier transforms, \( r(x) = f(x) \otimes f^*(x) \) and \( h(x) \otimes h^*(x) \) of the intensities of these patterns are shown in Fig. 3.18 together with \( [f(x) \otimes f^*(x) - h(x) \otimes h^*(x)] \). \( h(x) \) is clearly seen as that part of \([f(x) \otimes f^*(x) - h(x) \otimes h^*(x)]\) above the threshold \( V \) for \( x > 0 \). The reconstructed radiation pattern, \( H(u) \), is superimposed on Fig. 3.17. There is close agreement with the measured \(|H(u)|\).

### 3.4.4 H.P.M. Technique With Off Axis Point Source Reference

One further refinement is required to make the measurement techniques described in sections 3.4.2 and 3.4.3 completely practical for large antennas. The techniques as described require the test antenna to be mounted off-axis on the turntable. This may be impractical for large antennas, especially for antennas having their own steering systems. If the antenna is mounted at the centre of the turntable then the point source at \( x = 0 \) may still be used as a reference provided the three measurement technique described at the end of section 2.3.3 is used. Alternatively, the H.P.M. techniques
with complete or partial separation can be used if the point source is moved off the centre of the turntable.

As described in section 3.4.2 a point source at the centre of the turntable can be simulated using a fixed reference antenna fed by a signal of constant modulus and phase. A point source situated a distance \( T_R \) off-axis can be simulated by a fixed reference source of constant modulus but with its phase varying as \((-2\pi T_R u)\) radians. A continuously variable phase shifter inserted in the feed line of the reference antenna and controlled by the rotation of the turntable (for example by the output from a \( \sin \theta \) potentiometer on the turntable axis) could give this phase shift.

The following procedure may be used to obtain the aperture distribution of an antenna using an H.P.M. technique with a variable phase reference.

Procedure F - H.P.M. Technique With Off Axis Point Source Reference

(F.1) Mount the test antenna to be measured at the centre of a turntable.

(F.2) Set up a fixed reference antenna as described in step (D.2). Insert a continuously variable phase shifter into the feed line to the fixed reference antenna. Control the rotation of the phase shifter so that, as the turntable rotates, the phase shifter provides a phase shift of \(-2\pi T_R u\) radians, where \( T_R \) is a constant such that \( T_R > L_2/2 \).

(F.3) Measure \( |F(u)| \) the radiation pattern of the test antenna and fixed reference antenna with varying phase.

(F.4) If \( T_R > 3L_2/2 \), obtain \( h(x) \) using steps (D.4) and (D.5).
(F.5) If $3L_c/2 > T_R > L_c/2$, obtain $h(x)$ using steps (E.4) - (E.7).

Procedure F was tested experimentally using the arrangement shown in Fig. 3.19. The antenna to be tested, a 3.9 cm diameter circular transducer, was mounted at the centre of the turntable. This transducer, shown in Fig. 3.21, had a narrow (3 mm width) strip of acoustically absorbing sponge rubber placed over its aperture at right angles to the x axis to introduce an error into $h(x)$ at a known value of $x = -0.55$ cm. The sinθ output required to control the phase shifter was obtained by feeding the output from a linear potentiometer on the turntable shaft into a diode function generator. The phase shifter is described in Appendix 1. For the measurement, $T_R$ was chosen to be -7.0 cm, giving an effective separation between the point source reference and $h(x)$ of 5.15 cm. Thus, the H.P.M. method with complete separation is applicable. The phase of the input signal to the fixed reference antenna was varied as $-14\pi \sin \theta$ radians. $|F(u)|$, the pattern of the combination of the variable phase reference and the test antenna is shown in Fig. 3.20. $r(x)$, the F.T. of $|F(u)|^2$, is shown in Fig. 3.21 and because the effective separation between the point reference source and $h(x)$ is greater than the extent of $h(x)$, the unknown aperture distribution can be obtained directly by inspection of Fig. 3.21. The position of the perturbation in $h(x)$ is accurately indicated. Fig. 3.22 shows the actual measured pattern, $H(u)$, of the test antenna. The small differences in sidelobe levels on opposite sides of the main beam are caused by the perturbation in $h(x)$. Also shown in Fig. 3.22 is the
reconstructed $H(u)$ obtained as the F.T. of $h(x)$. There is good agreement between measured and reconstructed values of $H(u)$.

In all the H.P.M. techniques described in previous sections, it has been necessary that the test antenna be rotated about an axis contained in its aperture plane. An added advantage in using a reference signal whose phase is controlled by the rotation of the test antenna, as is the case in Procedure $F$ above, is that the method can be simply adapted to antennas having their own steering mechanisms which often do not rotate about axes contained in their aperture planes. Suppose that an antenna is rotated about an axis situated at a distance $z_0$ behind its aperture plane (for example see the antenna on the turntable in Fig. 4.3). Denote by $H_0(u)$ the radiation pattern of the antenna measured about this axis. If $H(u)$ is the radiation pattern measured about the centre of the aperture it can be simply shown (Brown and Jull, 1961) that

$$H_0(u) = H(u) \exp(j2\pi z_0 \cos \theta / \lambda) \quad (3.10)$$

In Procedure $F$ above (step $F.2$) the phase of reference source was controlled by the rotation of the turntable to give a phase shift of $-2\pi T_R u$ radians, where $u = \sin \theta / \lambda$. Now control the phase of the reference to give a phase shift of $-2\pi (T_R \sin \theta - z_0 \cos \theta) / \lambda$ radians. That is

$$G(u) = \exp(-j2\pi (T_R \sin \theta - z_0 \cos \theta) / \lambda) \quad (3.11)$$

Recall (section 2.3.1) that when an H.P.M. technique is used to measure an antenna, the required information is obtained from
the cross product term $H(u)G^*(u)$. If $G(u)$ is given by (3.11), then (3.10) shows that

$$H_0(u)G^*(u) = H(u) \exp(j2\pi R_\nu u) \quad (3.12)$$

so that the phase shift proportional to $z_0 \cos \theta$ has been removed allowing standard H.P.M. reconstruction to be carried out even though the test antenna has not been rotated about the centre of its aperture. The crux of the method is the provision of the phase shifted reference specified by (3.11). Write (3.11) as

$$G(u) = \exp(-j2\pi T \sin \theta' / \lambda) \quad (3.13)$$

where

$$T = (R_2^2 + z_0^2)^{\frac{1}{2}}, \quad \theta' = \theta + \tan^{-1}(-R_2/z_0) \quad (3.14)$$

In Procedure F, the phase of the reference varied in proportion to $\sin \theta$. This could have been done by controlling a phase shifter with the output from a $\sin \theta$ potentiometer mounted on the turntable shaft. (3.13) shows that we now require a phase shift proportional to $\sin \theta'$, where $\theta'$ is defined in (3.14). This can be accomplished by simply rotating the potentiometer on the shaft through $\tan^{-1}(-R_2/z_0)$ radians, so that it gives an output proportional to $\sin \theta'$ instead of $\sin \theta$. This is the only change to Procedure F that is needed.

3.4.5 Discussion of H.P.M. Techniques With Point Source

References

The simplest of the H.P.M. techniques with point source reference is certainly the complete separation method described in section 3.4.2. The only difference between this
technique and a normal radiation pattern measurement is that the antenna must be positioned off axis on the turntable and a constant reference signal radiated. The subtraction technique of section 3.4.3 is only slightly more complex and has the advantage of not requiring an extended antenna range. The necessity for the subtraction step in the computations can be expected, in general, to reduce the accuracy of this method. The variable phase reference method of section 3.4.4 will only be needed when the antenna cannot be placed off axis on the turntable. The double subtraction method described at the end of section 2.3.3, although likely to be error sensitive because of the subtraction steps, may be useful if the provision of a continuously variable phase shift is not convenient.

All the methods have the advantage that no physical link is required between opposite ends of the antenna range. If such a link is available, for instance in inside laboratories, the methods can still be useful. In this case, the fixed reference antenna may be replaced by electronically injected reference signals. The signal from the receiving antenna and the reference signal from the link would be summed at the detector. This procedure will often be more convenient than direct phase measurement using a phase comparator (Knott, 1969).
FIGURE 3.1. EXPERIMENTAL ARRANGEMENT

FIGURE 3.2. MEASURED RADIATION PATTERNS FOR EXAMPLE OF SECTION 3.2.1.
Figure 3.3. $r(x)$ (modulus only shown) for example of Section 3.2.1.

Figure 3.4. Radiation pattern of test antenna for Section 3.2.1.
FIGURE 3.5. COMPUTED APERTURE DISTRIBUTION OF TEST ANTENNA FOR SECTION 3.2.1.

FIGURE 3.6. RADIATION PATTERNS FOR SECTION 3.2.2.
FIGURE 3.7. RECONSTRUCTED $h(x)$ FOR SECTION 3.2.2

FIGURE 3.8. ANTENNA ARRAY AND EQUIVALENT APERTURE DISTRIBUTION
FIGURE 3.9. POSITIONS AND RELATIVE EXCITATIONS OF ELEMENTS IN ACOUSTIC ARRAY

NORMALIZED MODULUS OF PATTERN

FIGURE 3.10a ARRAY RADIATION PATTERNS
FIGURE 3.10b. OUTSIDE PARTS OF FOURIER TRANSFORMS OF $|F_A(u)|^2$ AND $|F_B(u)|^2$

- F.T. OF $|F_A(u)|^2$  --- MOD,  --- PHASE
- F.T. OF $|F_B(u)|^2$  ××× MOD,  ○○○ PHASE

FIGURE 3.10c. F.T. OF $|F(u)|^2$−$|H(u)|^2$ FOR 7 ELEMENT ARRAY
FIGURE 3.11. EQUIVALENT APERTURE DISTRIBUTION FOR CIRCULAR ARRAY

FIGURE 3.12. MEASURED RADIATION PATTERN FOR FIRST EXAMPLE OF SECTION 3.4.2.
FIGURE 3.13. $r(x)$ THE F.T. OF $|F(u)|^2$ FOR FIRST EXAMPLE OF SECTION 3.4.2. (MODULUS ONLY)

FIGURE 3.14. $|H(u)|$ FOR FIRST EXAMPLE OF SECTION 3.4.2.
Figure 3.15. Aperture distribution for 2nd example of Section 3.4.2.

Figure 3.16. |\(H(\omega)\)| for second example of Section 3.4.2.
FIGURE 3.17. RADIATION PATTERNS FOR EXAMPLE OF SECTION 3.4.3.

FIGURE 3.18. FOURIER TRANSFORMS OF RADIATION PATTERNS FOR EXAMPLE OF SECTION 3.4.3.
FIGURE 3.19. EXPERIMENTAL ARRANGEMENT FOR EXAMPLE OF SECTION 3.4.4.

FIGURE 3.20. MEASURED PATTERN, $|F(u)|$, FOR EXAMPLE OF SECTION 3.4.4.
FIGURE 3.21. \( r(x) \) FOR EXAMPLE OF SECTION 3.4.4.

FIGURE 3.22. \( H(y) \) FOR EXAMPLE OF SECTION 3.4.4.

--- MEASURED MODULUS, --- MEASURED PHASE
○○○ COMPUTED MODULUS AND PHASE
CHAPTER 4. Extensions to H.P.M. Techniques

In this Chapter some extensions to the H.P.M. technique described in Chapters 2 and 3 are described.

4.1 Use of a Reference Signal With Phase Shift Linear With Angle

In section 3.4.4 it was shown that the H.P.M. technique using a point reference source can be used when the test antenna is mounted at the centre of the turntable, provided the fixed reference antenna is fed by a signal of constant modulus and phase varying as \((-2\pi T_R u)\) radians. The choice of \(T_R\) determines whether the complete separation or partial separation H.P.M. techniques are appropriate. The required phase shift is a linear function of \(u\) and therefore must be varied linearly with \(s\). Practically, it would be more convenient to use a phase shift proportional to \(\theta\). In this case the phase shifter could be mechanically geared directly to the turntable shift. The case in which the fixed reference antenna is fed with a signal of constant modulus, and phase varying as \(2\pi T_R \theta/\lambda\) is now considered. In this case \(G(u)\), the radiation pattern of the reference source is given by

\[
G(u) = \exp(jkT_R \theta) \tag{4.1}
\]

where \(k = 2\pi/\lambda\). To determine whether the \(G(u)\) given in (4.1) is suitable to be the reference signal in a H.P.M. technique, it is necessary to examine \(g(x)\), the equivalent source distribution that produces \(G(u)\). \(g(x)\) is the F.T. of \(G(u)\).

\[
g(x) = \int_{-\infty}^{\infty} \exp(jkT_R \theta) \text{rect}(u\lambda/2) \exp(-j2\pi ux) \, du \tag{4.2}
\]
where \( \text{rect}(u\lambda/2) \) has been included because radiation patterns can only be measured for \(|u| < 1/\lambda\). Recall that \( u = \sin \theta / \lambda \), so that (4.2) can be written

\[
\begin{align*}
G(x) &= \frac{k}{4\pi} \int_{-\pi/2}^{\pi/2} \exp[jk(T_R \theta - x \sin \theta) + j\theta] d\theta \\
&\quad + \frac{k}{4\pi} \int_{-\pi/2}^{\pi/2} \exp[jk(T_R \theta - x \sin \theta) - j\theta] d\theta \\
&= \frac{k}{4\pi} \int_{-\pi/2}^{\pi/2} \exp[jk(T_R \theta - x \sin \theta)] d\theta \tag{4.3}
\end{align*}
\]

If \( kT_R \) or \( kx \) is large the integrand of the integrals in (4.3) is highly oscillatory so that the principle of stationary phase (Jones 1964, pg. 449) may be used to examine the behaviour of \( G(x) \) qualitatively. The stationary phase points of the integrand of (4.3) occur when

\[
\frac{d}{d\theta} (T_R \theta - x \sin \theta \pm \theta/k) = 0,
\]

where the plus and minus signs refer respectively to the first and second integrals in (4.3). The stationary phase points occur when

\[
\cos \theta = \frac{T_R \pm 1/k}{x} \tag{4.4}
\]

Since the range of integration is \(|\theta| < \pi/2\), \( \cos \theta \) satisfies \( 0 < \cos \theta < 1 \). Therefore, (4.4) shows that, for \( kT_R \) or \( kx \) large, \( G(x) \) will only have significant value for \( x > T_R - 1/k \).

Suppose now that \( G(x) \) is to be used as the reference source for an H.P.M. measurement to determine an unknown source distribution \( h(x) \) of width \( L_2 \) centred at the origin.

Recall, from section 2.3.1, that in an H.P.M. method the quantity obtained from the measurement is \( r(x) \), where \( r(x) \) is defined in (2.13). Consider the terms on the R.H.S. of (2.13). (4.2) shows that \( |G(u)|^2 = 1 \), so that \( G(x)G^*(x) \) is a delta function and can be neglected. If \( kT_R \) is large, \( G(x) \) is small for \( x < T_R - 1/k \), so that the cross-
correlation term $h(x)\phi g^*(x)$ will only be significant for $x > (T_R - 1/k - L_2/2)$. Thus, if $kT_R$ is chosen large enough, $h(x)\phi g^*(x)$ can be displaced far enough from the origin to enable it to be removed from $r(x)$ by inspection. This shows that the $G(u)$ in (4.1) is a suitable reference signal for an H.P.M. method.

Let us examine the behaviour of $g(x)$ further. Integrating (4.3) by parts gives

$$g(x) = \frac{-\sin[k(T_R^* - 2 - x)]}{kx} + \frac{kT_R}{2\pi x} \int_{-\pi/2}^{\pi/2} \exp(jkT_R\theta - jkx\sin\theta) d\theta$$

(4.5)

Consider the integral in (4.5) under the conditions

$$kx >> 1, \quad kT_R >> 1, \quad (x - T_R) >> 1$$

(4.6)

The integrand is highly oscillatory and the principal of stationary phase shows that the integral only has significant value in the region of $\cos\theta = T_R/x$. Then the error in extending the limits of the integral to $-\pi$ and $\pi$ is small. In this case, if $T_R$ is chosen such that $kT_R$ is an integer then the integral takes the standard integral form for a Bessel function (Abramowitz and Stegun 1965, pg. 360). Under condition (4.6) the first term on the R.H.S. of (4.5) is negligible compared to the second so that

$$g(x) \approx kT_R J_{kT_R}(kx)/x$$

(4.7)

(4.7) indicates an important feature of the behaviour of $g(x)$. (4.7) shows that $g(x)$ is not of finite extent and, from the asymptotic form for Bessel functions with large arguments (Abramowitz and Stegun 1965, Pg. 364), $g(x)$
oscillates with a period of $\lambda$ when $x \gg T_R$. The behaviour of $g(x)$, examined above on a qualitative basis, can be verified quantitatively by accurately evaluating the integral in (4.2). The simplest way to do this is to treat the integral as a Fourier transform and evaluate it using the Fast Fourier Transform algorithm, with $G(u)$ sampled at a sufficiently small interval to make aliasing errors negligible (see section 8.1). As an example $g(x)$, computed using the F.F.T. algorithm, is shown in Fig. 4.1 for the particular case $T_R = 20$ wavelengths. The two important features of $g(x)$ that have been predicted above, the negligible amplitude for $x < T_R$ and the rapidly oscillating behaviour for $x \gg T_R$, are clearly seen.

Although the extent of $g(x)$ is infinite, the effective extent of $h(x)\overline{g^*}(x)$ (i.e. the range of $x$ for which $h(x)\overline{g^*}(x)$ is above some noise threshold) need not be very large because of the oscillatory nature of $g(x)$. If $h(x)$ is a smooth function, adjacent lobes of $g(x)$ will tend to cancel and only that part of $g(x)$ in the region of $x \approx T_R$ will contribute to $h(x)\overline{g^*}(x)$. An example will illustrate this point. Fig. 4.2 shows $r(x)$, computed as the F.T. of $|F(u)|^2$ using the F.F.T. algorithm, for two cases. In both cases $G(u)$ is given by (4.1) with $T_R = 20$ wavelengths, corresponding to the $g(x)$ shown in Fig. 4.1. In both cases $H(u)$ corresponds to an $h(x) 10$ wavelengths wide centred at the origin. In one case $h(x)$ is a uniform distribution and in the other it is a cosine distribution. The autocorrelation of the reference, $g(x)\overline{g^*}(x)$, has been subtracted out of Fig. (4.2). Because $h(x)$ and $g(x)$ are effectively separated by a distance of
approximately 15 wavelengths, the H.P.M. technique with complete separation is applicable and \( h(x) \otimes g^*(x) \) can be obtained by inspecting \( r(x) \). The expected behaviour of \( h(x) \otimes g^*(x) \) is evident in Fig. 4.2. When \( h(x) \) is a uniform distribution, so that it has sharp discontinuities at its edges, \( h(x) \otimes g^*(x) \) has large effective extent. When \( h(x) \) is a smooth cosine distribution, \( h(x) \otimes g^*(x) \) is more compact. In an actual measurement the outer low level parts of \( h(x) \otimes g^*(x) \) will be lost in measurement noise and Fig. 4.2 shows that the lost information will be much less for smooth distributions. Notice that the decorrelation step required to obtain \( h(x) \) from \( h(x) \otimes g^*(x) \) is straightforward because \( G(u) \) has constant modulus.

The following H.P.M. procedure can be used to obtain an unknown aperture distribution \( h(x) \) using a reference signal whose phase varies linearly with \( \theta \).

**Procedure 6 Use of Reference Signal With Phase Proportional to \( \theta \).**

(G.1) Mount the test antenna to be measured at the centre of the turntable.

(G.2) Repeat step (F.2) except that the phase of the reference should be \( 2\pi T_R \theta \) where \( T_R > L_2 + 1/k \).

(G.3) Measure \( |F(u)| \) as in step (F.3).

(G.4) Compute \( r(x) \) using \( |F(u)|^2 \rightarrow r(x) \). Obtain \( e(x) \) as \( e(x) = r(x), x > T_R - 1/k \).

(G.5) Compute \( E^*(u) \) using \( e(x) \rightarrow E(u) \).

(G.6) Compute \( H(u) = E(u) \exp(-jkT_R \theta) \).

(G.7) Compute \( h(x) \) using \( H(u) \rightarrow h(x) \).
4.2 A H.P.M. Technique Analogous to Gabor In-Line Holography

Most of the H.P.M. techniques described above rely on the reference and unknown sources being laterally separated. The amount of the separation between the two sources determines which particular H.P.M. technique is applicable.

Conventional holography in its original form (Gabor, 1948) did not require this lateral separation between sources. The reference and unknown sources were "in-line". The sources were, however, separated with respect to distance from the hologram plane. An H.P.M. technique analogous to this type of holography is now described.

Consider the experimental setup shown in Fig. 4.3. Instead of being laterally offset from the centre of the turntable, the test antenna has its aperture plane displaced a distance $z_{0}$ from the x axis, as shown. Such a situation would apply, for instance, to a large antenna with its own steering mechanism when the pivot point is situated behind the aperture. In this case the pattern of the test antenna is $H(u) \exp(j2\pi z_{0}\cos\theta/\lambda)$ (Brown and Jull, 1961). If the fixed reference antenna provides a signal at the receiving antenna such that $G(u) = S$, where $S$ is a constant, then

$$|F(u)|^2 = H(u)H^*(u) + SS^* + S^*H(u)\exp(j2\pi z_{0}\cos\theta)$$

$$+ SH^*(u)\exp(-j2\pi z_{0}\cos\theta)$$

(4.12)

If $|S| \gg |H(u)|$, the term $H(u)H^*(u)$ may be neglected in (4.12), otherwise $|H(u)|$, the modulus of the radiation pattern of the test antenna by itself must be measured and $|H(u)|^2$ subtracted from $|F(u)|^2$. The constant term $SS^*$ may be
removed by inspection. Define $E(u)$ to be the sum of the last two terms in (4.12), which may be isolated as just shown. So

$$E(u) = S^*H(u) \exp(j2\pi z_0 \cos \theta) + SH^*(u) \exp(-j2\pi z_0 \cos \theta) \tag{4.13}$$

which shows that $E(u)$ is the sum of the radiations from a distribution $h(x)$ situated at $z = z_0$ and a distribution $h^*(-x)$ situated at $z = -z_0$. Now "focus" attention on the $z = z_0$ plane by multiplying $E(u)$ by $\exp(-j2\pi z_0 \cos \theta)$. Define $\Gamma(u)$ as

$$\Gamma(u) = E(u) \exp(-j2\pi z_0 \cos \theta) \tag{4.14}$$

Then, defining $\gamma(x)$ to be the F.T. of $\Gamma(u)$, (4.13) and (4.14) show that

$$\gamma(x) = h(x) + \int_{-\infty}^{\infty} H^*(u) \exp(-j4\pi z_0 \cos \theta) \exp(-j2\pi u x) du \tag{4.15}$$

Thus, $\gamma(x)$ consists of the required source distribution, $h(x)$, plus an unwanted term which corresponds to the radiation from a source $h^*(-x)$ situated at $z = -z_0$.

The integral in (4.15) is simply the field distribution at a distance $2z_0$ in front of an aperture distribution $h^*(-x)$. With a conventional Gabor "in-line" hologram, when a plane wave is used as a reference signal, $z_0$ is effectively infinite and the distorting term in (4.15) is negligible. However, it is unlikely that this will ever be the case for the measurement shown in Fig. 4.3, where $z_0$ will usually be smaller than $L_2$. The strength of the near field of an antenna is appreciably less than the field strength in the aperture only at distances greater than $L_2^2/\lambda$ (Hansen 1964, Pg. 38). For example, Fig. 4.4 shows the field at various distances $z_0$ in
front of an aperture of width $L_2 = 10\lambda$ where $h(x)$ is a cosine distribution. The distributions in Fig. 4.4 were obtained by computing the F.T. of $H(u) \exp(j2\pi z_0 \cos \theta / \lambda)$ using the F.F.T. algorithm. Fig. 4.4 demonstrates that the distorting term in (4.15) is certainly not negligible for $z_0$ of the same order as $L_2$. The only case in which this procedure could be of practical use is when the input to the fixed reference antenna is phase shifted as $\exp(-j2\pi z_0 \cos \theta)$. In this case, $z_0$ can be made as large as desired and

$$h(x) = \lim_{z_0 \to \infty} \gamma(x) \quad (4.16)$$

4.3 Use of H.P.M. Techniques in the Near Field

All of the H.P.M. techniques described so far have been suitable for use only in the far field. Some near field applications of the H.P.M. theory are now described.

4.3.1 Measurement of Narrow Beam Antennas

As mentioned in section 1.4, one reason for measuring the phase of an antenna's radiation pattern is to enable its far-field radiation pattern to be predicted from a measured near-field pattern. This problem is now examined for narrow beam antennas.

Consider a one-dimensional aperture distribution (section 1.2.1) $h(x)$, of width $L_2$ centred at the origin. Define $V(R, \theta)$ to be the field measured in the Fresnel region of the antenna at a distance $R$ and an angle $\theta$. $\theta$ is the angle measured from the direction perpendicular to the $x$ axis. Then for any angle in the range $|\theta| < 90^\circ$, it can be shown (Hansen 1964, Pg. 26) that a suitable expression for $V(R, \theta)$ is
\[ V(R, \theta) = \exp(-j\theta) \frac{\exp(-jkR)}{R} \int_{-\frac{L_2}{2}}^{\frac{L_2}{2}} h(x) \exp[jk(x\sin\theta - \frac{x^2}{2R}(1-\sin^2\theta))] \, dx \] 

(4.17)

The phase shift introduced by the last term in the argument of the exponential in (4.17) is \( x^2\sin^2\theta/2\lambda R \) wavelengths and the maximum value of this phase shift is \( L_2^2\sin^2\theta/8\lambda R \) wavelengths.

Suppose that a preliminary measurement shows that \( V(R, \theta) \) is negligible for all \( \theta \) greater than some value, call it \( \theta_{\text{max}} \). This means, for example, that if the noise level for the pattern measurement is \( N \), then

\[ V(R, \theta) < N, \quad |\theta| > \theta_{\text{max}} \] 

(4.18)

If \( V(R, \theta) \) is negligible for \( |\theta| > \theta_{\text{max}} \), then the maximum value of the phase shift introduced by the last term in the argument of the exponential in (4.17) is \( L_2^2\sin^2\theta_{\text{max}}/8\lambda R \). Assuming, as is usual (Hansen 1964, Chap. 1), that phase shifts less than \( \lambda/16 \) may be neglected, the phase shift \( L_2^2\sin^2\theta_{\text{max}}/8\lambda R \) may be neglected provided

\[ R > \frac{2L_2^2\sin^2\theta_{\text{max}}}{\lambda} \] 

(4.19)

Notice that if \( \theta_{\text{max}} = 90^\circ \), this reduces to the usual far field condition \( R > 2D^2/\lambda \). Define \( H_R(u) \) to be the radiation pattern measured at distance \( R \).

\[ V(R, \theta) = \frac{\exp(-j\theta)}{R} H_R(u) \] 

(4.20)

Then, assuming that condition (4.19) holds (4.20) and (4.17) shows that

\[ H_R(u) = \int_{-\frac{L_2}{2}}^{\frac{L_2}{2}} h(x) \exp(-j\pi x^2/\lambda R) \exp(j2\pi ux) \, dx \] 

(4.21)
where $u = \sin \theta / \lambda$. Define a modified aperture distribution $h_R(x)$ as

$$h_R(x) = h(x) \exp(-j \pi x^2 / AR) \quad (4.22)$$

(4.21) and (4.22) show that $H_R(u)$ and $h_R(x)$ are a F.T. pair satisfying

$$h_R(x) \rightarrow H_R(u), \quad H_R(u) \rightarrow h_R(x) \quad (4.23)$$

If $H_R(u)$ is measured in modulus and phase, $h_R(x)$ can be computed using (4.23). $h(x)$ can be determined using (4.22) which then leads to $H(u)$ using $h(x) \rightarrow H(u)$.

To summarize, the following procedure can be used to obtain the aperture distribution and far-field radiation pattern of an antenna from a measured near field pattern.

Procedure H - Fourier Transform Method for Obtaining Far Field From Measured Near Field Radiation Pattern

(H.1) At any distance $R$ in the Fresnel region of the antenna, carry out a preliminary measurement of $|H_R(u)|$ to determine the maximum angle, $\theta_{\text{max}}$, for which $|H_R(u)|$ is significant. If the inequality (4.19) is satisfied proceed to step (H.2). If (4.19) is not satisfied, $R$ must be increased and step (H.1) repeated.

(H.2) Measure $H_R(u)$ in modulus and phase. Compute $h_R(x)$ using $H_R(u) \rightarrow h_R(x)$.

(H.3) Compute $h(x)$, the aperture distribution of the antenna using $h(x) = h_R(x) \exp(j \pi x^2 / AR)$.

(H.4) Compute $H(u)$ using $h(x) \rightarrow H(u)$.

Notice that for antennas with narrow main beams whose radiation patterns are effectively confined to a finite part
of the u axis, condition (4.19) can be satisfied at distances considerably smaller than the usual \( R = 2Lc^2/\lambda \) criterion. For example, if \( H_R(u) \) is negligible for \(|\theta| > 30^\circ\), (4.23) is satisfied at a quarter of the usual far-field distance whilst if \( \theta_{\text{max}} \approx 10^\circ \), (4.23) holds at distances greater than only 3% of the usual far-field distance.

Procedure H is based on the same theory as the method used by Bickmore (1957) to measure the far-field radiation pattern of an antenna in the Fresnel region of the antenna. Both Procedure H and Bickmore's method require that the \( \sin^2 \theta \) term in the exponent in (4.17) is negligible. This assumption is known as the small-angle Fresnel approximation (Hansen 1964, p.27) and is also basic to the near-field – far-field technique proposed by Bates and Elliott (1956). In Procedure H the far-field pattern is obtained by introducing a quadratic phase shift into the aperture distribution of the antenna (step H.3). Bickmore (1957) introduces this quadratic phase shift into the aperture distribution of an antenna by physically deforming the antenna into a slight circular arc. For an array, when the antenna cannot be physically deformed the required phase shift can be obtained by changing the excitations of the individual elements (Hansen 1964, p.42). The method of simply inserting the required phase shift computationally, as is proposed in Procedure H, does not appear to have been previously suggested. The availability of the F.F.T. algorithm for rapid computation of Fourier transforms makes the computations required in Procedure H straightforward.
The chief disadvantage of Procedure H is that it is only applicable in measurement situations that are effectively two dimensional. Thus, Procedure H could be used to measure in the near field of a linear antenna or array or, for the case of an antenna having a two dimensional aperture, at a range in the near field of one aperture dimension and in the far field of the other. A further disadvantage of Procedure H is that it assumes that the antenna used to measure the pattern of the test antenna is a point source. It would be difficult to take the finite size of a measuring antenna into account using this simple F.T. near field - far field method.

As an example of the application of the F.T. method, Fig. 4.5 shows the result of processing a near-field measurement made by Jensen (1970). The antenna measured was a slotted waveguide array 21.8 cm long ($L_2 = 21.8$ cm) and 1 cm high radiating at a wavelength $\lambda = 3.407$ cm. Radiation patterns were measured in the horizontal plane. The far field radiation pattern measured at 8 m ($R = 5.7 \frac{L_2^2}{\lambda}$) is shown in Fig. 4.5, as is the near field pattern measured at 51.8 cm ($R = 0.37 \frac{L_2^2}{\lambda}$). The phase of the near field pattern, which is not shown in Fig. 4.5, was also measured. If steps (H.2) - (H.4) are to be exactly applicable for $R = 51.8$ cm, condition (4.19) requires that the near field pattern should be negligible for $\theta$ greater than $\theta_{\text{max}} = 25^0$. Even though Fig. 4.5 shows that this requirement is not satisfied, the far-field pattern, computed from the near-field using steps (H.2) - (H.4), which is shown in Fig. 4.5, shows very good agreement with the measured far-field pattern for the main beam and first few sidelobes. This demonstrates the usefulness of
Procedure H and suggests that it will often be applicable at
distances shorter than the distance required to satisfy
condition (4.19).

Since Procedure H is based on a F.T. relationship between
the near-field pattern and the antenna aperture distribution,
H.P.M. techniques may be used to obtain the phase information
required in step (H.2). Suppose that the test antenna is
being measured at a distance $R$ in its near field, and that
step (H.1) shows that the near field F.T. relationship is
applicable. Fig. 4.6 shows the antenna with a point reference
source of strength $S$ situated at its edge. At the receiving
antenna the reference source provides a signal $G(u)$ where

$$G(u) = S \exp\left[ jk\left( \frac{L_2}{2} \sin \theta - \frac{L_2^2}{8R} (1 - \sin^2 \theta) \right) \right] \quad (4.24)$$

Since condition (4.19) holds, the $\sin^2 \theta$ term in (4.23) may be
neglected and $G(u)$ can be expressed as $S' \exp(j\pi L_2 u)$, where
$S' = S \exp(-jkL_2^2/8R)$. Then the intensity of the measured
combined pattern, $|F(u)|^2$, is given by

$$|F(u)|^2 = |H_R(u)|^2 + |S|^2 + S' \exp(j\pi L_2 u)H_R^*(u)$$
$$+ S^* \exp(-j\pi L_2 u)H_R^2(u) \quad (4.25)$$

If $|H_R(u)|$ is measured as well as $|F(u)|$, then the F.T. of
$|F(u)|^2 - |H_R(u)|^2$ will give $h_R(x)$, (just as the example of
section 3.3.3 gave $h(x)$). $h_R(x)$ can then be processed as in
steps (H.3) and (H.4).

It has been shown in the previous paragraph that if an
experimental arrangement equivalent to Fig. 4.6 is used, the
antenna aperture distribution and far-field pattern can be
obtained from modulus-only measurements in the near field. The point source situated at the edge of the antenna can be provided by using a fixed reference antenna fed by a variable phase signal, as described in section 3.4.4. Notice, however, that the F.T. relationship established in (4.21) only holds if the antenna is rotated about its centre, so that any of the H.P.M. techniques that require the antenna to be rotated off axis cannot be applied unless the distance \( R \) is increased. For example, if the antenna is rotated about its edge, the pattern must be measured at a range \( 2R \).

4.3.2 Phase Correction of the Reference Signal

All of the H.P.M. techniques described so far, except for the special case of section 4.3.1, require that the measured radiation patterns be the F.T. of the radiating source distribution. This requires that the "hologram", \( |F(u)|^2 \), be measured in the far field of the combination of \( h(x) \) and \( g(x) \). If \( g(x) + h(x) \) has a total effective extent of \( L \), then a range of \( 2L^2/\lambda \) ensures that \( F(u) \) is the F.T. of \( g(x) + h(x) \) to within the accuracy of the usual \( \lambda/16 \) phase error. For those H.P.M. techniques for which \( L > L_2 \), the required increase in range beyond the \( 2L_2^2/\lambda \) far-field range of the test antenna will usually be inconvenient to provide. Consider the case shown in Fig. 4.7 in which the phase shifted reference signal from the fixed antenna provides an equivalent point reference source a distance \( T_R \) off-axis.

The pattern of a point source in this position is

\[
\exp \left[ jk(T_R \sin \theta - \frac{T_R^2}{2R_n} \cos^2 \theta) \right].
\]

Provided then that the input signal to the fixed reference antenna has its phase varied as \( k(T_R \sin \theta - \frac{T_R^2}{2R_n} \cos^2 \theta) \) or, neglecting the constant phase
term, as \( k(T_R \sin \theta + T_R^2 \frac{\partial}{\partial R} \sin^2 \theta) \), \(|F(u)|\) can be measured at
\( R = 2L_2^2/\lambda \) even though the effective extent of \( f(x) \) is
greater than \( L_2 \).

4.3.3 Angular Fourier Series

Up to this point, antenna patterns have been expressed
as functions of \( u = \sin \theta/\lambda \) in the range \(|\sin \theta| < 1\). We now
consider patterns as functions of \( \theta \) for all \( \theta \) in the range
\( 0 < \theta < 2\pi \). Consider the measurement set up shown in Fig.
4.8. Let the test antenna have a pattern \( B(\theta) \) for this value
of \( R_n \). Then, because it is a periodic function of \( \theta \) with
period \( 2\pi \), \( B(\theta) \) may be represented to any desired accuracy by
an angular exponential Fourier series.

\[
B(\theta) = \sum_{n=-N}^{N} b_n \exp(jn\theta)
\]  

(4.26)

where the \( b_n \) are complex numbers which vary with \( R \). The
integer \( N \) is determined by the accuracy required of the
representation. The \( b_n \) are given by

\[
b_n = \frac{1}{2\pi} \int_{0}^{2\pi} B(\theta) \exp(-jn\theta) \, d\theta
\]  

(4.27)

Unlike the Fourier transform relationships between the radia-
tion pattern, \( H(u) \), and the aperture distribution, \( h(x) \),
which is only valid in the far field, the Fourier series
relationship between the angular radiation pattern, \( B(\theta) \) and
the angular harmonic coefficients, \( b_n \), is valid at all ranges.
Provided the \( b_n \) are known in modulus and phase, \( B(\theta) \) can be
completely determined in modulus and phase using (4.26).

Fig. 4.9a shows a typical \( B(\theta) \) and its corresponding \( b_n \) (this
\( B(\theta) \) and \( b_n \) are for illustrative purposes only and are not
accurately computed). A method analogous to H.P.M. techniques
will now be described which allows the $b_n$ to be obtained from only the modulus of a measured angular radiation pattern.

Suppose that the phase shifter in Fig. 4.8 varies the phase of the reference signal as $M\theta$ radians, where $M$ is an integer. Then $A(\theta)$, the combined pattern of the reference and test antennas together is given by

$$A(\theta) = B(\theta) + \exp(jM\theta)$$  (4.28)

Consider the intensity $|A(\theta)|^2$ of this combined pattern. Since $|A(\theta)|^2$ is a periodic function of $\theta$ it can be represented as an angular exponential Fourier series

$$|A(\theta)|^2 = A(\theta)A^*(\theta) = \sum_{m=-L}^{L} a_m \exp(jm\theta)$$  (4.29)

where

$$a_m = \frac{1}{2\pi} \int_{0}^{2\pi} |A(\theta)|^2 \exp(-jm\theta) \, d\theta$$  (4.30)

and $L$ is an integer. Substituting (4.26) and (4.28) into (4.29) gives

$$|A(\theta)|^2 = \sum_{m=-L}^{L} a_m \exp(jm\theta) = 1 + \sum_{n=-N}^{N} |b_n|^2$$

$$+ \sum_{m=-2N}^{-1} \exp(jm\theta) \left( \sum_{n=-N-m}^{N} b_n^* b_{n+m} \right)$$

$$+ \sum_{m=1}^{2N} \exp(jm\theta) \left( \sum_{n=-N-m}^{N-m} b_n^* b_{n+m} \right) + \sum_{m=-M-N}^{-M+N} b_{M+m}^* \exp(jm\theta)$$

$$+ \sum_{m=M-N}^{M+N} b_m^* \exp(jm\theta).$$  (4.31)

(4.31) shows the relationship between the coefficients $a_m$, which can be computed straightforwardly from a measured $|A(\theta)|^2$ using (4.30), and the unknown coefficients $b_m$. The
highest order, \( L \), of the \( a_m \) coefficients is \( N+M \). (4.31) shows that when \( M > 3N \)

\[
a_m = 0, \quad M-N > |m| > 2N
\]  
\[
a_m = b_{M-m}^*, \quad m > M-N
\]  

so that, as shown in Fig. 4.9b, the \( b_m \) can be obtained by inspection from the \( a_m \). This method of determining the \( b_m \) is completely analogous to the H.P.M. technique, described in section 3.4.4, with a point source reference and complete separation. \( |A(\theta)|^2 \) and the \( a_m \) are analogous to \( |F(u)|^2 \) and \( r(x) \) respectively. The requirement that \( M > 3N \) is equivalent to the requirement that \( g(x) \) be separated from \( h(x) \) by a distance greater than \( L_2 \). In Fig. 4.9b, the \( a_m \) in the range \( |m| < 2N \) are analogous to the autocorrelation terms \( g(x) \oplus g^*(x) + h(x) \oplus h^*(x) \), and the \( a_m \) in the range \( |m| > M-N \) are analogous to the cross correlation terms \( g(x) \oplus h^*(x) \) and \( h(x) \oplus g^*(x) \). In analogy with the Partial Separation H.P.M. technique with a point source reference, if \( M \) is in the range \( 3N > M > N \) then the unknown \( b_m \) can be obtained by inspection of the angular harmonic coefficients of \( [|A(\theta)|^2 - |B(\theta)|^2] \).

An important consideration in the application of the method is the choice of \( N \) and \( M \). \( N \), the number of terms required in (4.26), can be estimated from the dimensions of the test antenna. Brown and Jull (1961) have shown that if the test antenna has an aperture length of \( L_2 \), then a suitable estimate of \( N \) is given by

\[ N \approx \pi L_2 / \lambda. \]
The phase shift required in the reference signal is $2\pi M$ radians per revolution of the turntable, varying linearly with $\theta$. It is important that $M$ be an integer since only in this case will the simple relationship between the $b_m$ and $a_m$ obtained by equating coefficients in (4.31) be valid. The required phase shift can be easily provided in practice using a continuously variable rotating phase shifter mechanically coupled to the turntable shaft through a gear box with a gear ratio of 1:M.

The coefficients $a_m$ of the measured intensity pattern $A(\theta)$ can be simply calculated from (4.30) using the F.F.T. algorithm. Since there are $2(M+N)+1$ coefficients, $A(\theta)$ must be sampled at at least this number of points equally spaced in the range $2\pi > \theta > 0$. As with the H.P.M. techniques (Section 2.4.4), a noise threshold must be applied to the $a_m$ to separate out the $b_m$. The level of this threshold gives an indication of the error in the $b_m$. A check on the accuracy of the $b_m$ can be obtained by comparing the $|B(\theta)|$ calculated from (4.26) with the measured $|B(\theta)|$.

The method described above can be used to obtain the modulus and phase of an antenna radiation pattern $B(\theta)$ at any distance from an antenna. Only the modulus of the combined pattern $A(\theta)$, defined by (4.27), need be measured. The method would be of particular value when the far-field radiation pattern is to be predicted from a near field measurement. Particular near field – far field techniques (Brown and Jull, 1961) require the $b_m$ coefficients as their starting points. The procedure to be followed in applying the technique described above is now given.
Procedure I - Measurement of Radiation Pattern Over $360^\circ$

(I.1) Mount the test antenna to be measured at the centre of a turntable.

(I.2) Set up a fixed reference antenna as described in step (D.2) (section 3.4.2). Insert a continuously variable phase shifter into the feed line of the fixed reference antenna. Control the rotation of the phase shifter so that, as the turntable rotates, the phase shifter provides a phase shift of $M\theta$ radians, where $M$ is an integer such that $M > \frac{3\pi L_2}{\lambda}$.

(I.3) Measure $|A(\theta)|$, the radiation pattern of the test antenna and fixed reference antenna with varying phase, over the range $0 < \theta < 2\pi$.

(I.4) Carry out a Fourier analysis of $|A(\theta)|^2$ using equation (4.30) to obtain the $a_m$ coefficients.

(I.5) Apply a noise level $V$ (section 2.4.4) to the $a_m$ and identify the coefficients $b_{M-m}^*$ as the $a_m$ in the range $m > \frac{2\pi L_2}{\lambda}$ that are above $V$.

(I.6) Reconstruct $B(\theta)$, the radiation pattern of the test antenna, by inserting these $b_m$ in equation (4.26).

The results of a practical application of Procedure I are now described. The antenna measured, which is shown in Fig. 4.10, was a parabolic cylinder reflector fed by a strip transducer. This antenna, the phase shifter necessary for the measurement and the computer controlled data sampling techniques were provided by final year undergraduate students in the Electrical Engineering Department of the University of Canterbury (Dickens, 1970; Dilks, 1970; Dunlop, 1970). The pattern of the antenna was measured in the $xz$ plane at a
distance of 56 cm. The width of the antenna, $L_2$, was 15 cm so that, with $\lambda = 1$ cm, (4.34) gives $N \approx 47$. $M$ was chosen to be 200 so that $M > 3N$. The patterns $|A(\theta)|$ and $B(\theta)$ are shown in Fig. 4.11. The coefficients $a_m$, computed using (4.30) are shown in Fig. 4.12. When the threshold $V$ is applied the unknown coefficients $b_m^*$ are obtained as those $a_m$ centred on $m = 200$. The $B(\theta)$ reconstructed using these $b_m$ in (4.26) with $N = 50$ agrees reasonably well with the actual measured $B(\theta)$, as shown in Fig. 4.11.

Finally in this section we mention the method of Knott (1969) who obtains the modulus and phase of a radiation pattern for $0 < \theta < 2\pi$. His method is analogous to the double subtraction H.P.M. technique that was described at the end of section 2.3.3 and has the advantage over Procedure I that the reference signal does not have to have continuously varying phase shift. For completeness, a procedure equivalent to Knott's (1969) method is included below.

**Procedure J - Pattern Measurement Over 360° Without Variable Phase Shift**

(J.1) Mount test antenna to be measured at the centre of a turntable.

(J.2) Set up a fixed reference antenna as described in Step (D.2).

(J.3) Measure $|B(\theta)|$, the radiation pattern of the test antenna.

(J.4) Measure $|A_1(\theta)|$, the radiation pattern of the test antenna and reference together. Call the constant reference signal $S$.

(J.5) Phase advance the signal to the reference antenna by
\[ \pi/2 \text{ radians. The reference signal is now } jS. \text{ Measure } |A_2(\theta)|, \text{ the pattern of the test antenna and reference together.} \]

\[(J.6) \text{ Compute } B(\theta) \text{ using}
\]
\[ B(\theta) = \frac{1}{2} [ |A_1(\theta)|^2 - |B(\theta)|^2 - |S|^2 + j(|A_2(\theta)|^2 - |B(\theta)|^2 - |S|^2)] \]

Notice that the calculation in step \((J.6)\) provides both the modulus and phase of \(B(\theta)\). Only if \(|B(\theta)|\) obtained in step \((J.6)\) agrees well with \(|B(\theta)|\) measured in step \((J.3)\) can the phase obtained in step \((J.6)\) be considered reliable. This provides a useful check on the accuracy of the measurement. Knott (1969) only computed the phase of \(B(\theta)\).

### 4.4 Advantages of H.P.M. Techniques in a Distorting Medium

It is well known that the holographic method of obtaining the image of an object offers distinct advantages when the medium is inhomogeneous and distorts waves propagating through it (Gaskill, 1968). The H.P.M. techniques offer similar advantages, as will now be demonstrated.

Fig. 4.13 shows a typical H.P.M. setup. Test and reference antennas with the same polarisation are illuminated by a radiation source situated in the far field of the combination of \(h(x) + g(x)\). As the angle \(\theta\) varies, it is possible in some measurement situations (for instance when a satellite is used as a source) that the source will maintain neither the same aspect towards nor the same distance from the two antennas. Furthermore, variations in the medium can distort the signal received by \(g(x)\) and \(h(x)\). Thus, the signal transmitted towards the antennas will vary virtually unpredictably. Let \(S(u)\) be the combined effect of the source variation and the distortion introduced by the propagation
medium. The signals received by the antennas, in combination and separately, are proportional not to \( F(u), G(u) \) and \( H(u) \) but rather to \( \hat{F}(u), \hat{G}(u) \) and \( \hat{H}(u) \) where

\[
\begin{align*}
\hat{F}(u) &= F(u)S(u) \\
\hat{G}(u) &= G(u)S(u) \\
\hat{H}(u) &= H(u)S(u)
\end{align*}
\] (4.35)

The measured hologram, \( |\hat{F}(u)|^2 \), can be expanded as

\[
|\hat{F}(u)|^2 = |G(u)S(u) + H(u)S(u)|^2 = |S(u)|^2(|G(u)|^2 + |H(u)|^2 + G(u)H^*(u) + H(u)G^*(u))
\] (4.36)

Notice that the phase of \( S(u) \) does not appear in (4.36).

This shows that phase distortion introduced by the propagation medium does not affect the accuracy of H.P.M. techniques. The assumption must be made of course that the test and reference antennas are sufficiently close together so that distortions in the propagation paths to both antennas are the same.

Since \( G(u) \), the radiation pattern of the reference antenna, is known \( |S(u)| \) can be obtained

\[
|S(u)| = \frac{|\hat{G}(u)|}{|G(u)|^2}
\] (4.37)

so that the undistorted hologram can be calculated

\[
|F(u)|^2 = |\hat{F}(u)|^2/|S(u)|^2
\] (4.38)

allowing \( H(u) \) and \( h(x) \) to be obtained using the method appropriate to the particular H.P.M. technique in use.
FIGURE 4.1. \( g(x) \) THE FOURIER TRANSFORM OF \( G(\omega) \). FOR \( x < 15 \),
\( g(x) \) IS SMALL (LESS THAN 0.01) FOR \( x > 45 \). PERIOD OF
OSCILLATIONS APPROACHES 1 WAVELENGTH,
ENVELOPE DECREASES MONOTONICALLY.

FIGURE 4.2. \( r(x) \) FOR A UNIFORM AND COSINE APERTURE DISTRIBUTION
--- \( h(x) \) UNIFORM, ----- \( h(x) \) COSINE
FIGURE 4.3. TEST SET UP WITH REFERENCE AND UNKNOWN SOURCES NOT LATERALLY SEPARATED.

FIGURE 4.4. FIELD DISTRIBUTION AT DISTANCE $2z_0$ IN FRONT OF A 10-WAVELENGTH APERTURE WITH COSINE ILLUMINATION.
FIGURE 4.5. NEAR AND FAR-FIELD RADIATION PATTERNS

--- MEASURED NEAR-FIELD, ---- MEASURED FAR-FIELD (JENSEN [1970])
○--○ FAR-FIELD PATTERN COMPUTED USING METHOD OF SECTION 4.3.1.

FIGURE 4.6. TEST ANTENNA WITH POINT REFERENCE SOURCE AT ITS EDGE.
**Figure 4.7.** Phase shifted reference signal in near-field.

**Figure 4.8.** Antenna pattern measured over range $0 < 	heta < 2\pi$. 

\[ \text{Effective position of point reference sources} \]

\[ \text{Signal source} \rightarrow \text{Phase shifter} \rightarrow \text{Fixed reference antenna} \]

\[ \text{Phase shift} \ k \left( \frac{R_n \sin \theta + R_n^2 \sin^2 \theta}{\partial R_n} \right) \]

\[ \text{Test antenna} \rightarrow \text{Fixed reference antenna} \]

\[ \text{Signal source} \rightarrow \text{Phase shift} \rightarrow \text{Fixed reference antenna} \]
FIGURE 4.9a A TYPICAL $b(\theta)$ AND ITS ANGULAR COEFFICIENTS $b_n$ (MODULUS ONLY SHOWN)

FIGURE 4.9b. COEFFICIENTS $a_m$ WHEN $M > 3N$

FIGURE 4.10. CYLINDRICAL PARABOLIC ANTENNA MEASURED IN SECTION 4.3.3.
Figure 4.11 Parts of $|A(\theta)|$ and $B(\theta)$ for example of section 4.3.3.

Figure 4.12 Angular spectrum coefficients for $|A(\theta)|^2$. 
FIGURE 4.13. H.P.M. TECHNIQUE IN PRESENCE OF DISTORTION.
CHAPTER 5. The Zeros of Entire Functions

In Chapters 2, 3 and 4 it has been demonstrated that in certain cases a source distribution can be obtained from only the modulus of its radiation pattern if part of the source distribution is known. This known part of the source distribution was called the reference source. In Chapters 5, 6 and 7 we consider the problem of reconstructing source distributions that do not contain a reference source and for which the phase of the radiation is not known at all or is incompletely or inaccurately known. In this chapter, we first establish the necessary theory for the reconstruction procedures which are discussed in Chapters 6 and 7, and then review the previous work in the field.

5.1 Entire Function Theory

The theory given in sections 5.1.1, 5.1.2 and 5.1.3 is based on the approach used by Bates (1969a, 1969b).

5.1.1 Fourier Transforms in the Complex Plane

The source reconstruction procedures described in the previous chapters have been based on the F.T. relationship between \( f(x) \) and \( F(u) \), where \( f(x) \) is the equivalent one dimensional source distribution of a coherent radiating source and \( F(u) \) is the far-field radiation pattern of the source. For the work of Chapters 5, 6 and 7 it is convenient to extend these definitions so that \( f(x) \) is a distribution of any physically important quantity and \( F(u) \), the F.T. of \( f(x) \), is some physically measurable quantity. It is desired to obtain \( f(x) \) by measuring \( F(u) \). Possible examples of an \( f(x) - F(u) \)
pair are: source distribution-radiation pattern for coherent sources, brightness temperature distribution - complex fringe visibility for incoherent sources (see section 6.1.1) and spectral distribution - complex fringe visibility for interference spectroscopy (Steel 1967, p.231). In previous chapters \( f(x) \) has been composed of two separate parts, \( g(x) \) and \( h(x) \). For Chapters 5, 6 or 7 \( f(x) \) will not in general be separable into two parts, although for particular cases discussed in Chapter 7 it is convenient to reintroduce \( g(x) \) and \( h(x) \). As in previous chapters (c.f. (2.5)), in Chapters 5, 6 and 7 it is assumed that \( f(x) \) is of finite extent.

When \( f(x) \) is of finite extent, it is useful to think of \( F(u) \) as being the value on the real axis of \( F(w) \) where \( w \) is the complex variable

\[
w = u + jv
\]  
(5.1)

Although \( F(w) \) can only be measured for \( w \) real, \( F(w) \) and \( f(x) \) are related through the Fourier transforms

\[
F(w) = \int_{-\beta}^{\beta} f(x) \exp(j2\pi wx) \, dx \tag{5.2}
\]

\[
f(x) = \int_{-\infty}^{\infty} F(u) \exp(-j2\pi ux) \, du \tag{5.3}
\]

where \( 2\beta \) is the extent of \( f(x) \) (in the notation of (2.5) \( 2\beta = L, A \) is assumed zero). The finite limits on the integral in (5.2) are important because Theorem 10 of Paley and Wiener (1934, p.13) states that the F.T. of a function of finite extent is an entire function of exponential type. An entire, or integral, function is one that is analytic everywhere in the finite complex plane. An entire function of exponential
type (integral function of order one) is an entire function whose behaviour as \(|w|\) tends to infinity is \(O(e^{C|z|})\), where \(C\) is a constant (Titchmarsh 1939, Chap. 8). An important property of entire functions is that their behaviour is determined by the positions of their zeros of which there are a denumerable infinity. Thus, Morse and Feshbach (1953, p. 385) shows that \(F(w)\) can be expressed as

\[
F(w) = \exp(A+Bw) \prod_{n=1}^{\infty} \left(1 - \frac{w}{a_n}\right) \exp\left(\frac{w}{a_n}\right)
\]

where \(A\) and \(B\) are constants and the set of complex numbers \(\{a_n\}\) is the set of zeros of \(F(w)\). A feature of \(|F(w)|\) that will be found useful later is that it falls smoothly to its zeros. This follows from the fact that an analytic function cannot have true maxima or minima. It can only have zeros or saddle points (Morse and Feshbach 1953, p. 369).

A simple example will illustrate some of the points made above. Fig. 5.1 shows a particular \(f(x)\) and Fig. 5.2 shows the \(F(u)\) corresponding to this \(f(x)\). \(F(w)\) has zeros in the complex plane at the points shown in Fig. 5.3. Notice that the behaviour of \(F(u)\) is strongly dependent on the positions of the members of \(\{a_n\}\); zeros close to the real axis have a large effect on \(F(u)\) whilst zeros further from the real axis have a reduced effect. Fig. 5.4 shows a section of \(F(w)\) in the complex plane. The importance of the zeros in determining the behaviour of \(F(w)\) is clearly evident. (The example given in Figs 5.1-5.4 corresponds to Example 1 of Chapter 6.)
5.1.2 Complex Zeros and the Sampling Theorem

Since \( f(x) \) is of finite extent \( 2\beta \), it can be represented to any desired accuracy by a Fourier series of the form

\[
f(x) = \sum_{n=-M}^{M} F_n \exp(-j\pi nx/\beta), \quad |x| < \beta
\]

where \( M \) is an integer that increases as the accuracy of the Fourier series representation increases and the \( F_n \) are complex constants. Substituting (5.5) into (5.2) gives

\[
F(w) = \frac{\sin(2\pi \beta w)}{\pi} \sum_{n=-M}^{M} \frac{(-1)^n F_n}{w - n/2\beta}
\]

where

\[
F_n = F(\frac{n}{2\beta})/2\beta
\]

(5.6) and (5.7) are simply statements of the well known Fourier sampling theorem which states that \( F(u) \), the Fourier transform of a function of finite extent \( 2\beta \), is completely determined by the samples of \( F(u) \) at intervals of \( 1/2\beta \).

(5.6) can be written in the form

\[
F(w) = P(w) Q(w)
\]

where

\[
Q(w) = \frac{\sin(2\pi \beta w)}{2\pi \beta} \prod_{m=-M}^{M} (w - \frac{m}{2\beta})
\]

and

\[
P(w) = \sum_{n=-M}^{M} (-1)^n P(\frac{n}{2\beta}) \prod_{m=-M}^{M} (w - \frac{m}{2\beta})
\]

where \( \prod(n) \) denotes a continued product with the term \( m = n \) missed out. All the zeros of \( Q(w) \) are real and the product expansion for \( \sin(w) \) (Abramowitz and Stegun 1965, p.75) shows that these are equally spaced along the real axes at the points \( u = n/2\beta \), where \( n \) is an integer and \( |n| > M \).
Inspection of (5.10) shows that $P(w)$ is a polynomial of order $2M$, so that

$$P(w) = c \prod_{n=1}^{2M} (w-w_n)$$  \hspace{1cm} (5.11)

where the constant $c$ is given by

$$c = F(0)(-1)^M \prod_{m=1}^{M} \left( \frac{w_m}{w} \right)^2 \prod_{m=1}^{2M} w_n$$  \hspace{1cm} (5.12)

The set $\{w_n = u_n + jv_n\}$ of order $2M$ is the set of zeros of $P(w)$, where $P(w)$ is defined by (5.10). Since all the zeros of $Q(w)$ are real, $\{w_n\}$ is the set of complex zeros of $F(w)$ although, for any particular $F(w)$, any number of the $w_n$ can be real.

The set $\{a_n\}$ mentioned in section 5.1.1 is composed of the zeros of $P(w)$ and the zeros of $Q(w)$.

Equations (5.6) and (5.7) show that $F(w)$ is determined completely by its value at its sample points and defining the zeros $\{w_n\}$ can be regarded as simply an alternative method of storing the information contained in these samples. The zeros of $P(w)$ contain all the information about the $2M$ non-zero samples of $F(u)$, whilst the zeros of $Q(w)$ ensure that, as a result of the finite limits on the summation in (5.5), $F(u)$ is zero at all other sample points. Although the concept of representing $F(u)$ in terms of its complex zeros is by no means as familiar as the usual sampling theorem, in many cases the zeros provide a more useful representation.

5.1.3 The Source Distributions Defined by $|F(u)|$

It is well known that in general there are many possible source distributions all of which will give rise to the same modulus of radiation pattern. Using the complex zeros of $F(w)$ it is very simple to obtain all the possible source distributions that have the same $|F(u)|$, as will now be demonstrated.
If any of the complex \( w_n \) in (5.11) are replaced by their complex conjugates, \(|F(u)|\) is not changed and so \(|F(u)|\) remains unchanged although \(\arg F(u)\) is altered. It is convenient to use the notation of Bates (1971) to define all the possible \(F(u)\) that can be formed by this process of zero "flipping". Let \( L \) of the \( 2M \) members of \( \{w_n\} \) be complex. Define

\[
\tau_p = 0 \text{ or } 1
\]

where \( p \) is a positive integer. There are \( 2^L \) distinct sequences \( \{\tau_p\} \) if \( 1 < p < L \). Define

\[
F^{(m)}(w) = \tau_m(w)F(w)
\]

(5.14)

where

\[
\tau_m(w) = \prod_{p=1}^{L} \frac{w - \tau_p w_p}{w - \tau_p w_p^*}
\]

(5.15)

where the \( \tau_p \) in (5.15) are taken from the \( m \)th of the \( 2^L \) sequences \( \{\tau_p\} \). In (5.14), \( \tau_m(w) \) replaces the \( w_p \) by \( w_p^* \) if \( \tau_p = 1 \) and leaves the \( w_p \) unchanged if \( \tau_p = 0 \). It follows from (5.8), (5.11) and (5.14) that

\[
|F^{(m)}(u)| = |F(u)|
\]

(5.16)

Define \( f^{(m)}(x) \) using (5.3) with \( f(x) \) and \( F(u) \) replaced by \( f^{(m)}(x) \) and \( F^{(m)}(u) \). Then the \( f^{(m)}(x) \) are source distributions that all have the same \(|F(u)|\). Sheehan (1962), Walther (1963) and Bates (1969b) show that these \( f^{(m)}(x) \) are in fact the only source distributions of finite extent that can have the same \(|F(u)|\).

An important consequence of the discussion given above is that, given only \(|F(u)|\), it is possible to determine all the \( f^{(m)}(x) \). \( R(u) \), the intensity of \( F(u) \), is defined by
(2.1). \( r(x) \), the F.T. of \( R(u) \) defined in (2.12), is the autocorrelation of \( f(x) \), and is necessarily of extent 48. From the property of the finite F.T. discussed in section 5.1.1, \( R(u) \) can also be continued throughout the complex plane as an entire function of exponential type. From (2.1), the continuation of \( R(u) \) is given by

\[
R(w) = F(w)F^*(w^*)
\]  

(5.17)

\( R(w) \) is characterized by the positions of its zeros in the complex plane. Substituting (5.8) into (5.17) gives

\[
R(w) = P(w)P^*(w^*)Q(w)Q^*(w^*)
\]  

(5.18)

(5.18) shows that the real zeros of \( Q(w) \) are double real zeros of \( R(w) \). From (5.18), \(|P(u)|^2\) can be obtained as

\[
|P(u)|^2 = R(u)/|Q(u)|^2
\]  

(5.19)

Equation (5.11) shows that \(|P(u)|^2\) is a polynomial of order \( 4M \) whose zeros are given by the set \( \{w_n\} \cup \{w_n^*\} \). Thus the complex zeros of \( R(w) \) are positioned symmetrically about the real axis. Define the set \( \{\zeta_m = p_m + jq_m, q_m > 0\} \) with 2M members to be the zeros of \( P(w)P^*(w^*) \) situated in the upper-half-plane (U.H.P.). Any real zeros of \( P(w)P^*(w^*) \) must be double real zeros; only one of these is included in \( \{\zeta_m\} \).

From the \(|P(u)|^2\) defined in (5.19) the members of \( \{\zeta_m\} \) can be computed, but it will not be known, for any particular \( \zeta_m \), whether \( \zeta_m \) or \( \zeta_m^* \) is a member of \( \{w_n\} \). Define \( F^{(0)}(w) \) to be the \( F(w) \) that has all its zeros in the U.H.P. Then \( F^{(0)}(w) \) can be obtained using (5.8), (5.9) and (5.11) with \( F(w) \) and \( w_m \) replaced by \( F^{(0)}(w) \) and \( \zeta_m \) respectively. All the possible \( F^{(m)}(w) \) can be obtained from (5.14) and (5.15) with \( F(w) \).
replaced by \( F^{(0)}(w) \). Define \( f^{(m)}(x) \), \( 1 < m < 2M \), to be all the possible distributions that give rise to \( R(u) \). Then the \( f^{(m)}(x) \) can be obtained using \( F^{(m)}(u) \rightarrow f^{(m)}(x) \).

An important question to be answered is: in what cases is knowledge of only \( |F(u)| \) sufficient to allow unambiguous reconstruction of \( f(x) \). The most important case is when \( f(x) \) is Hermitian symmetric, \( f(x) = f^*(-x) \) (a real symmetric function satisfies this requirement). In this case, Theorem 3 of section 5.3 shows that the zeros of \( F(w) \) are either on the real axis or are positioned symmetrically about it so that the zeros of \( F(w) \) can be chosen unambiguously from the zeros of \( R(w) \). A case of less practical importance is when \( f(x) \) is known to be a function that is non-increasing or non-decreasing. In this case Theorem 8 of section 5.3 shows that the zeros of \( F(w) \) are on only one side of the real axis so that the zeros of \( F(w) \) can be chosen unambiguously from the zeros of \( R(w) \).

5.2 Review of Literature on Applications of Complex Zero Theory

The theory of the Fourier Transforms in the complex plane and the properties of entire functions are well established mathematically (Paley and Wiener, 1934; Titchmarsh, 1939). The zeros of entire functions have been studied in detail (Nardan, 1949; Boas, 1954; Levin, 1964). In this section previous applications of these theories to various types of radiation problems are reviewed. Where necessary the nomenclature of sections 5.1.1-5.1.3 is used to describe results from the papers discussed. Some of the more useful results
concerning the complex zeros of entire functions are listed in section 5.3.

Taylor (1955), in his classic paper on antenna radiation pattern synthesis, was one of the first to apply entire function theory to antenna pattern problems. He shows, using similar theory to sections 5.1.1 and 5.1.2, that an antenna pattern is characterized by the positions of its zeros in the complex plane. In particular, he is interested in the behaviour of the zeros as \(|u|\) becomes large. His main result is to show that the positions of the zeros for large \(|u|\) is determined completely by the behaviour of \(f(x)\) at its edges. Only patterns with zeros on the real axis are considered in detail. Hyneman and Johnson (1967) extend Taylor's results to the design of complex shaped-beam radiation patterns by displacing zeros of the radiation pattern from the real axis. They note that the power pattern is unchanged if any of the complex zeros are changed for their complex conjugates. Schell (1963) also proposes the zero "flipping" procedure of section 5.1.3 as a means of determining all the possible line source distributions that give a particular power pattern.

A problem similar to that of determining a source distribution by Fourier transformation of its measured radiation pattern is that of determining the distribution of energy in the spectrum of a light beam from the F.T. of the visibility of the interference fringes measured with a Michelson interferometer. Wolf (1962) shows that for assymetrical spectra, when both the modulus and phase of the fringe visibility are necessary but only the modulus can be measured conveniently, the analyticity of the fringe visibility and examination of
its zeros in the complex plane can be of value. The problem is slightly different from the one discussed in section 5.1.3 in that the spectral density function is not of finite extent so that the fringe visibility is not necessarily an entire function. However, Wolf shows that the fringe visibility is analytic in one half of the complex plane and, provided the positions of the zeros of the fringe visibility in that half plane are known, the phase of the visibility can be obtained from its modulus through a Hilbert transform relationship. Wolf considers the special case in which it is known a priori that there are no zeros in the half plane. In this case the phase of the visibility is uniquely determined by the modulus. Wolf is unable to find conditions to ensure zeros on only one side of the real axis, but illustrates the case with an accompanying example (Kano and Wolf, 1962). Roman and Marathay (1963) examine the same problem considered by Wolf (1962). Their attempts to place constraints on the positions of the complex zeros give only the relatively simple results that if \( f(x) \) is real the zeros of \( F(w) \) are on, or are symmetrically positioned about, the imaginary axis. If \( f(x) \) is real and positive, the possibility of zeros on the imaginary axis is removed. They suggest other ways in which the positive definiteness of \( f(x) \) could possibly be used to constrain the zeros of \( F(w) \) but are unable to obtain useful results. Works by Dialetis (1967) and Nussenzveig (1967) also deal with Wolf's (1962) problem. Nussenzveig gives a useful list of theorems that summarizes many of the properties of zeros of entire functions (see section 5.3). In particular, several of these theorems use the positive definiteness of
f(x) to place quite strong restrictions on the zeros of F(w).

In a mathematical work Sheehan (1962) uses the theory of entire functions and their zeros to establish several useful theorems concerning the autocorrelation and crosscorrelation of functions of finite extent. He examines the problem of obtaining all possible functions that have the same autocorrelation function and derives a zero "flipping" result similar to the method of section 5.1.3. Walther (1963) examines the same problem related to obtaining the phase of optical images formed by lenses with finite apertures. He formulates the procedure of section 5.1.3. Both Sheehan and Walther note that the zero flipping technique does not have a simple extension to two dimensional problems.

Several authors have applied entire function theory to the analysis and synthesis of signals. Bond and Cahn (1958) show that the zeros of a band limited function provide a useful alternative to the usual Fourier sampling theorem, as was discussed in section 5.1.2. The work of Hofstetter (1964) is an application of the zero flipping technique to signal synthesis problems. Hofstetter's paper contains several useful examples of the application of the technique. Others have used entire function theory to characterize bandlimited signals (Voelcker, 1966a, 1966b; Seckey, 1970; Fawc, 1970).

Finally in this section we consider the contributions of Altman, Bates and Fawle (1964) and Bates (1965, 1969a, 1969b, 1971). Many of the results presented in Chapters 6 and 7 are based on Bates' work. Altman, Bates and Fawle (1964) use entire function theory to attack the problem of determining
the phase of a scattering pattern from a measurement of only the modulus of the pattern. Like Wolf (1962), they note that the phase of scattering patterns which have their zeros confined to one side of the real axis can be obtained directly from the modulus using a Hilbert transform relationship. The only generally applicable constraint that they obtain is that if \( f(x) \) is even and complex, \( F(w) \) necessarily has zeros in both upper and lower half planes. To apply the theory of sections 5.1.2 and 5.1.3, it is necessary to know \( 2\beta \), the width of \( f(x) \). The most obvious way of obtaining an estimate of \( \beta \) when only \( |F(u)| \) is known, is to recognise that \( R(u) \), the F.T. of \( |F(u)|^2 \), is of width \( 4\beta \). Bates (1965) proposes an alternative to this method based on the Hilbert transform relationship mentioned above. In a detailed treatment of the subject Bates (1969a, 1969b) suggests the application of the theory outlined in sections 5.1.2 and 5.1.3 to the problem of reconstructing an astronomical brightness temperature distribution when only the intensity of the F.T. of the distribution is measured with an intensity interferometer. He gives a detailed computational procedure for determining all the possible distributions. Important practical considerations such as the effect of measurement noise and the truncation of the interferogram are included in the procedure. A suggestion is also made for using the zero techniques to correct phase errors in compound interferometers. In a later work (Bates, 1971) the application of this reconstruction technique to the problem of obtaining the aperture distribution of an antenna from only the modulus of its radiation pattern is proposed.
5.3 Some Useful Theorems on Zeros of Entire Functions of Exponential Type

A list of useful theorems concerning the zeros of entire functions of exponential type is now given. The theorems are stated without proof. The proofs (or the source of the proof) can be found in the references quoted with each theorem. Statements of the theorems are given in terms of the zeros of $F(w)$ where $F(w)$ and $f(x)$ satisfy (5.2) and (5.3). The theorems are grouped according to the constraints placed on $f(x)$.

**$f(x)$ complex and symmetrical**

**Theorem 1** (Altman Bates and Fowle, 1964): If $f(x)$ is symmetrical, $f(x) = f(-x)$, then the zeros of $F(w)$ are symmetrical through the origin of the complex plane. Thus, if $w_n$ is a zero of $F(w)$, $-w_n$ is also a zero.

**Theorem 2** (Altman Bates and Fowle, 1964): If $f(x)$ is symmetrical and complex, $F(w)$ must have complex zeros. Hence, from Theorem 1, $F(w)$ must have zeros in both the upper and lower halves of the complex plane.

**Theorem 3** (Hofstetter, 1964): If $f(x)$ is Hermitian symmetric, $f(x) = f^*(-x)$, then the zeros of $F(w)$ are either on the real axis or are positioned symmetrically about the real axis. Thus, if $w_n$ is a zero of $F(w)$, $w_n^*$ is also a zero.

**Theorem 4** (Hofstetter, 1964): If $f(x)$ is Hermitian symmetric and $F(u) > 0$ (this implies that $f(x)$ is an autocorrelation function), any real zeros of $F(w)$ must be of even multiplicity.
f(x) real

Theorem 5 (Roman and Marathay, 1963): If f(x) is real, 
\[ f(x) = f^*(x), \]
then the zeros of F(w) are either on the imaginary axis or are positioned symmetrically about the imaginary axis. Thus, if \( w_n \) is a zero of F(w), \( -w_n^* \) is also a zero.

Theorem 6 From Theorems 3 and 5, if f(x) is real and symmetric 
\[ f(x) = f^*(x) = f^*(-x), \]
the zeros of F(w) are positioned symmetrically about the real and imaginary axis. Thus, if \( w_n \) is a zero of F(w), \( -w_n^* \), \( w_n^* \) and \( -w_n^* \) are also zeros.

f(x) real and positive

Theorem 7 (Roman and Marathay, 1963): If f(x) is real and positive, \( f(x) > 0 \), F(w) cannot have zeros on the imaginary axis.

Theorem 8 (Nussenzveig, 1967): If f(x) is real and positive and if f(x) is a non increasing (non decreasing) function, then all zeros of F(w) have negative (positive) imaginary parts, except possibly when f(x) is piecewise constant.

Theorem 9 (Nussenzveig, 1967): If f(x) is real, positive, continuous and (except perhaps at a finite number of points) differentiable and if
\[ a < \frac{f'(x)}{f(x)} < b, \quad |x| < \beta \]
where a and b are constants, then all the zeros \( \{w_n = u_n + jv_n\} \) of F(w) lie within the open strip \( a > v_n > b \).

When \( \frac{f'(x)}{f(x)} \) is constant (for example when f(x) is a truncated exponential) the strip reduces to a line on which all zeros are situated.
Asymptotic behaviour of zeros

Theorem 10 (Nussenzveig, 1967): The average density of the zeros of $F(w)$ is equal to the width of $f(x)$. Thus, if $n(u)$ is the number of zeros of $F(w)$ with modulus less than $u$, then

$$\lim_{u \to \infty} \frac{n(u)}{2u} = 2\beta.$$ 

Theorem 11 (Taylor, 1955): The asymptotic behaviour of the zeros of $F(w)$ for large values of $u$ is determined by the behaviour of $f(x)$ at its edges. For $f(x)$ symmetric assume that

$$f(x) \sim K (\beta+x)^{\alpha} \quad \text{as} \quad x \to -\beta$$

$$f(x) \sim K (\beta-x)^{\alpha} \quad \text{as} \quad x \to \beta$$

where $\sim$ denotes "is asymptotic to", and $K$ and $\alpha$ are constants. Let the zeros of $F(w)$ be grouped in pairs $iw_1, iw_2$ etc., and numbered in such a way that the moduli of the members of successive pairs do not decrease. Then the members of the $n^{th}$ zero pair tend to the positions $w_n = i(\frac{n}{2\beta} + \frac{\alpha}{4\beta})$ as $n$ tends to infinity.

Theorem 12 (Nussenzveig, 1967): If $f(x)$ is continuous for $|x| < \beta$ and $f(\pm\beta) \neq 0$, and if $f'(x)$ is integrable, the zeros of $F(w)$ are asymptotically given by

$$w_n \approx \frac{n}{2\beta} - \frac{\alpha}{4\pi \beta} \ln \left[ \frac{f(-\beta)}{f(\beta)} \right] + \epsilon_n$$

where $\epsilon_n \to 0$ as $n \to \infty$. 
FIGURE 5.1. BRIGHTNESS TEMPERATURE FOR EXAMPLE OF SECTION 5.1.1.

FIGURE 5.2. INTERFEROGRAM $F(u)$ FOR THE $F(x)$ SHOWN IN FIG 5.1.

FIGURE 5.3. ZEROS OF $F(w)$ CORRESPONDING TO $F(x)$ IN FIG. 5.1.
Figure 5.4 Section of $F(w)$ corresponding to $f(x)$ shown in Fig. 5.1.
CHAPTER 6. Application of Complex Zero Techniques to Astronomical Interferometry

In this Chapter, the theory outlined in Chapter 5 is applied to the problem of obtaining phase information in Astronomical Interferometry.

6.1 Astronomical Interferometry

The requirement in radio astronomy for increased resolution capabilities of telescopes, and the difficulty of constructing single instruments with apertures large enough to provide this resolution, has led to the widespread use of the interferometer to obtain brightness temperature maps of astronomical radio sources. To a lesser extent interferometers have also found applications in optical astronomy. In sections 6.1.1 to 6.1.3 we consider the problem of obtaining phase information in astronomical interferometry and in the remaining sections of this Chapter develop solutions to the problem.

6.1.1 The Fourier Transform Relationship in Radio Interferometry

Interferometry at radio wavelengths was first used to obtain information about astronomical radio sources in 1946 (McCready et al., 1947, Ryle and Vonberg, 1946). Since this early work, interferometry has been very widely used in radio astronomy. Many detailed reviews of the field are available, including those of Ostrowsky (1965), Swenson and Mathur (1968) and Cohen (1969).
Detailed descriptions of the theory of the interferometer are also available (Swenson & Mathur, 1968), and the theory will not be repeated here. The starting point for the present work is the well known fact that an interferometer measures the Fourier transform of the brightness temperature distribution under observation.

Fig. 6.1a shows a schematic diagram suitable for any two-element interferometer such as a simple Michelson interferometer or a phase switched interferometer (Jennison 1966) or a correlation interferometer (Swenson & Mathur, 1968). The interferometer elements are separated by a distance d. f(x) is the equivalent one dimensional brightness temperature distribution of the part of the celestial sphere under observation, where \( x = \sin \theta \), and for sources of small angular extent centred on the zenith, \( x \approx \theta \). Examination of the fringes produced by the interferometer gives \( F(u) \), where

\[
u = d \cos \phi / \lambda \tag{6.1}
\]

In (6.1), \( \lambda \) is the wavelength corresponding to the centre frequency of the interferometer receiver. The modulus and phase of \( F(u) \) are given respectively by the normalized modulus and relative position of the fringes. \( F(u) \) is usually known as "complex fringe visibility" (Bracewell, 1961) but for simplicity, in the present work it will be referred to as the "interferogram". It can be shown (Swenson and Mathur, 1968) that \( f(x) \) and \( F(u) \) are a F.T. pair provided that the following assumptions are valid:

(a) The source \( f(x) \) is incoherent.

(b) The interferometer system has identical antennas and a
rectangular passband centred on frequency $f_0 = c/\lambda$ of width $\Delta f$. The receiving system is quasi-monochromatic, $\Delta f/f_0 \ll 1$.

(c) The brightness distribution and antenna patterns have insignificant variations in the passband.

(d) The interferometer interference pattern has a maximum in the direction $\phi$. This requires that the lengths of the transmission paths from the two antennas to the detector differ by a delay of $dsin\phi/c$, where $c$ is the velocity of light.

(e) The source is isolated and small in extent compared with the antenna pattern.

In the work reported in this chapter and the next, it is assumed that these conditions are satisfied so that a F.T. relationship exists between $f(x)$ and $F(u)$. If the interferometer elements are themselves high gain aerials (often necessary if sensitive measurements are to be made) then the field of view of the interferometer is limited to that part of the celestial sphere from which signals can be detected. Consequently, the field of view increases with the intensity of sources being observed and the post detector integration time; and it decreases with deterioration of receiver sensitivity and improvements (i.e. narrower main beams and lower sidelobe levels close to the main beams) in the radiation patterns of the aerials. Even though the field of view of the interferometer is a function of the total apparatus being used and of the sources being observed, it is nevertheless effectively finite. The sources will often be of smaller extent than the field of view. It is reasonable, therefore, to assume that
the observed brightness temperature distribution is of finite extent. Call this extent \( \theta \). Then \( f(x) \) and \( F(u) \) satisfy (5.2) and (5.3).

6.1.2 The Problem of Phase Measurement in Radio Interferometry

The theory of section 5.1.3 shows that, except in special cases, both the modulus and phase of the interferogram are required to determine the brightness temperature distribution unambiguously. In general, the modulus of the interferogram can be obtained with much better accuracy than the phase (Jennison 1966, Clark 1970). The chief factors causing difficulties in the measurement of the phase of the interferogram are instability in the atmosphere (Baars, 1967; Basart et al., 1970; Barton, 1971), variations in the links joining the interferometer elements and variable phase shifts in the electronic equipment and local oscillators (Jennison 1958 and 1966, Swenson and Mathur 1968). Multiple element (compound) interferometers (Jennison 1958, MacPhie 1966) have been introduced to reduce the difficulty of maintaining phase stability over large distances between the interferometer elements, but with these techniques measurement errors tend to accumulate (Jennison, 1958; Bates, 1969a). With the introduction of very-long-baseline interferometry, all the difficulties mentioned above are being accentuated (Cohen, 1969).

Finally, it should be mentioned that the problem of obtaining phase information is greatest when a source is observed by an intensity interferometer (post detection correlation interferometer) (Hanbury Brown and Twiss, 1954). In this case only \( R(u) \), the intensity of the interferogram is measured and no information at all is obtained about the
phase of the interferogram. This type of instrument has found only limited application in radio astronomy (Jennison and Das Gupta, 1956), mainly because of its insensitivity to weak sources.

6.1.3 Interferometry in Optical Astronomy

The phase sensitive interferometer whose application in radio astronomy is discussed above was first used as an optical instrument by Michelson (1920) who used it to determine the angular diameter of some large stars. Michelson measured only the modulus of the interferogram and useful measurements of the interferogram phase are to date non-existent, a consequence of the very severe practical difficulties encountered at optical wavelengths (Goodman, 1970a). These difficulties are chiefly the problems of maintaining equal path lengths in the two arms of the instrument and the quite considerable effects of atmospheric turbulence. Recent developments in optical interferometry (Goodman, 1970a) suggest that phase information may soon be available.

The baselines required to measure the diameters of any but the very largest stars are too great for the Michelson interferometer. The intensity interferometer of Hanbury Brown and Twiss (1954) has been used to provide the necessary resolution (Hanbury Brown et al., 1967a, 1967b). As with the radio intensity interferometer, the stellar intensity interferometer is severely limited in its sensitivity and no information about the phase of the interferogram is obtained.
6.2 Identification of Errors in a Measured Interferogram

In sections 6.3.1-6.3.3 some new methods are presented for removing phase errors from measured interferograms. As a preliminary to this, sections 6.2.3 and 6.2.4 describe methods for identifying phase errors in an interferogram. Some preliminary definitions are given in section 6.2.1 and section 6.2.2 describes three examples that will be used to illustrate the error identification procedures and the error removal techniques.

6.2.1 Preliminary Definitions

In the interferogram processing techniques described below, three distinct types of physical quantity are considered. True values are unsubscripted. Thus, using the nomenclature of section 6.1.1, \( f(x) \) is the actual (real and non-negative) brightness temperature on the part of the celestial sphere being observed and \( F(u) \) is the ideal interferogram (complete and uncorrupted by noise or distortion). Any measured interferogram is corrupted by noise and by distortion in the receiver and in the propagation path between the radio stars and the interferometer. Also, a measured interferogram must be incomplete because it is only possible to use a finite range of element spacings.

A subscript \( m \) denotes a measured quantity or a quantity that is computed directly from measured data. Thus, \( F_m(u) \) is the measured (incomplete and corrupted by noise and distortion) interferogram and \( f_m(x) \) is the brightness temperature distribution computed directly from it, using \( F_m(u) \rightarrow f_m(x) \), where (5.3) has been used and the arrow indicates a Fourier transform as defined in the glossary.
The extent of \( f(x) \) is \( 2\beta \) so that the extent of \( f_m(x) \) would be denoted by \( 2\beta_m \). A subscript \( r \) denotes a quantity obtained from some refinement procedure carried out on measured data to decrease the effects of measurement errors. Thus, \( F_r(u) \) and \( \beta_r \) are better estimates of \( F(u) \) and \( \beta \) than are \( F_m(u) \) and \( \beta_m \).

As mentioned in the previous paragraph, imperfections in measured interferograms are "truncation distortion" due to the finite maximum baseline and "measurement noise" which represents the combined effects of true noise and distortion in the receiver. It is convenient to separate measurement noise into two parts: modulus noise \( N(u) \) and phase noise \( \Psi(u) \). All imperfections are covered by the term "measurement distortion". When an interferogram is measured over a range of baselines up to \( u_{\text{max}} \), write

\[
|F_m(u)| = |F(u)| \text{rect}(u/2u_{\text{max}}) + N(u) \quad (6.2)
\]

\[
\arg F_m(u) = (\arg F(u)) \text{rect}(u/2u_{\text{max}}) + \Psi(u) \quad (6.3)
\]

where the rectangular gate function, \( \text{rect}(u) \), is defined in the glossary.

In section 6.3.2 a procedure is presented for removing phase errors from interferograms when the modulus of the interferogram is known more accurately than the phase. A criterion for comparing the relative importance of modulus and phase errors is required. Refer to Fig. 6.1b which shows the relationship between quantities occurring in (6.2) and (6.3) for a particular value of \( u \). OA represents the accurate quantity \( F(u) \), whilst OC represents the actual measured quantity \( F_m(u) \) which is contaminated by modulus and phase.
errors. The relative importance of the modulus error, $N(u)$, and the phase error, $\Psi(u)$, can be estimated by comparing the lengths $|BC|$ and $|BA|$ respectively. The phase error can be considered larger than the modulus error if

$$\frac{|BA|/|F(u)|}{|BC|/|F(u)|} > 1$$  \hspace{1cm} (6.4)

Examination of Fig. 6.1b shows that (6.4) holds provided $\Psi(u) > N(u)/|F(u)|$. Thus for any particular measured interferogram, we will say that the modulus information is more reliable than the phase information provided

$$\bar{\Psi(u)} > \left| \frac{N(u)}{F(u)} \right|$$  \hspace{1cm} (6.5)

where the superbar denotes an average value with respect to $u$. It is interesting to note that phase errors are typically in the range 4 to 40 degrees (Baars, 1967) so that $\Psi(u)$ is in the range 7% to 70%. Uncertainty in the modulus of the interferogram will usually be less than this.

The errors of truncation distortion and measurement noise in a measured interferogram correspond to the errors of truncation and noise in measured antenna radiation patterns that were discussed in sections 2.4.2 and 2.4.3. The discussion in those sections applies equally whether $F(u)$ is an antenna radiation pattern or an interferogram, and the expressions given can be used to estimate the error caused by measurement noise or truncation. Even though observed brightness temperature distributions can be considered to have finite extent, as explained in section 2.4.4 it will be necessary in practice to apply a threshold level to $r_m(x)$ or $f_m(x)$ to give them finite widths. Denote by $V_r$ and $V_f$ the
threshold levels for \( r_m(x) \) and \( f_m(x) \) respectively. For a computed distribution \( f_m(x) \), define \( x_1 \) and \( x_2 \) respectively as the least and greatest values of \( x \) for which \( |f_m(x)| > V_r \). Then the effective extent of \( f_m(x) \) is defined to be \( x_2 - x_1 \).

The use of \( V_r \) and \( V_f \) is illustrated in the examples given in sections 6.2.3 and 6.4.3. The choice of \( V_r \) is important since this determines the effective extent of \( r_m(x) \) which is used in sections 6.2.3, 6.3.2 and 6.4.3 to give the best available estimate of \( 4\beta \). Bates (1969a) suggests that \( V_r \) can be estimated using (2.38) but, as mentioned in section 2.4.2, this sometimes leads to a value for \( V_r \) that is excessively large and hence an estimate of \( \beta \) that is smaller than the correct value. If \( f(x) \) is reconstructed using a value of \( \beta \) that is too small, the interferogram will be undersampled and the resulting aliasing errors (section 8.1) in \( f(x) \) may be large. Reconstruction of \( f(x) \) with a value of \( \beta \) that is larger than necessary corresponds to oversampling of the interferogram which does not prevent a correct reconstruction, although it increases the amount of noise present in the reconstructed \( f(x) \). So, care must be taken to avoid too large an estimate for \( \beta \). The estimate of \( V_r \) used to obtain the results reported in this chapter was chosen so as to reject any negative parts in \( r_m(x) \) which gives the effective extent of \( r_m(x) \) as the width of its central lobe between zero crossings. This estimate of \( 4\beta \), if in error, is likely to be slightly larger than the correct value. In cases where it is found that the reconstructed \( f(x) \) has an effective extent somewhat smaller than that estimated from \( r_m(x) \) in this way (the reconstruction of Example 4 in section 6.4.3 is an
example of this), the reconstruction procedure can be repeated with a smaller, more accurate value of $\beta$. The value of $V_r$ chosen should not be so large that all negativeness in $f_m(x)$ is rejected since, if the negativeness is due to phase errors in $F_m(u)$, the reconstruction procedure of section 6.3.2 requires that it be included in the effective extent of $f_m(x)$. In a particular case, whether the negative parts are worth being concerned about is a matter for the astronomer's judgement and is affected by whether he thinks he could do any better. If a computed quantity $s(x)$ lies above, or below, a prescribed limit $W$ for intervals of what we consider (mainly subjectively) to be "significant extent" then we write

$$s(x) > W \text{ or } s(x) < W, \text{ } x \text{ of significant extent.}$$

Truncation distortion sometimes has negligible effect on the effective extents of computed distributions. However, truncation distortion always limits the detail that can be resolved in computed distributions. As explained in section 2.4.2, truncating the interferogram at $u = u_{\text{max}}$ wavelengths is equivalent to limiting the resolution in the $x$-domain to approximately $1/2u_{\text{max}}$, which is here defined as the "resolution limit" of the measurement.

6.2.2 Three Examples

The interferogram processing techniques described in sections 6.2.3, 6.2.4 and 6.3.2 are illustrated with the following three examples.

Example 1

This is an ideal case, the data being generated within the computer. $f(x)$ is taken as the sum of three half-wave cosines:
which, when substituted into (5.2), gives

\[ F(u) = \frac{4}{\pi} \sum_{k=1}^{3} \frac{A_k \beta_k \cos(2\pi \beta_k u) \exp(j2\pi x_k u)}{(1-(4\beta_k u)^2)} \]  

(6.7)

Fig. 6.2 shows \( f(x) \) and \( F(u) \) for the values of \( A_k, \beta_k \) and \( x_k \) given in Table 6.1.

<table>
<thead>
<tr>
<th>( k )</th>
<th>( A_k )</th>
<th>( \beta_k )</th>
<th>( x_k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0</td>
<td>3.0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>3</td>
<td>0.75</td>
<td>0.75</td>
<td>-4.5</td>
</tr>
</tbody>
</table>

Table 6.1: Values of \( A_k, \beta_k \) and \( x_k \) for Example 1. \( \beta_k \) and \( x_k \) are expressed in minutes of arc.

The "measured" interferogram \( F_m(u) \) is obtained from \( F(u) \) by making the following assumptions about the "measurement". The modulus of the interferogram is measured accurately out to a maximum baseline length, \( u_{\text{max}} \), of 1400 wavelengths. No measurements are made for longer baselines. The measured phase is distorted by the addition of a slowly varying sinusoid, a type of distortion which, if uncorrected, introduces large errors in the brightness temperature distribution, as Fig. 6.2a shows. So

\[ \psi(u) = 1.05 \sin(2\pi u/933) \]  

(6.8)

The phase \( \psi(u) \) is shown in Fig. 6.2a and \( f_m(x) \), obtained using \( F_m(u) \rightarrow f_m(x) \), is shown in Fig. 6.2a.
Example 2

This is an interferometer measurement of Taurus A, the radio source in the Crab Nebula, made at 1427 Mhz by Twiss, Carter and Little (1960). A three element interferometer of the type described by Jennison (1958) was used for the measurement. Fig. 6.3 shows the measured interferogram \( F_m(u) \), and \( f_m(x) \) obtained using \( F_m(u) \rightarrow f_m(x) \).

Example 3

This is a combination of two interferometer measurements made on radio source PO349-27 at 1403 Mhz by Ekers (1968), who did not measure the phase, and at 1425 Mhz by Fomalont (1967, 1968) who measured both the modulus and phase of the interferogram. It is assumed that the frequencies are sufficiently close to make \( f(x) \) the same for both measurements. Fig. 6.4b shows the measured interferogram \( F_m(u) \) which is taken as the average of the two measurements. \( f_m(x) \), obtained using \( F_m(u) \rightarrow f_m(x) \), is shown in Fig. 6.4a.

6.2.3 Comparison of Extents of Distributions

Comparison of the effective extents of \( f_m(x) \) and \( r_m(x) \) is the basis of a simple test for the presence of errors in a measured interferogram. The theoretical basis of the test, the details of which are presented below, is given in Appendix 2.

Procedure K

(K.1) Compute \( r_m(x) \), using \( |F_m(u)|^2 = R_m(u) \rightarrow r_m(x) \), and note its effective extent \( 4\beta_r \).

(K.2) Compute \( f_m(x) \), using \( F_m(u) \rightarrow f_m(x) \), and note its effective extent \( 2\beta_m \).

(K.3) Compute \( \gamma = 2u_{\max} |\beta - \beta_r| \). If \( |\gamma| < 1 \), arg \( F_m(u) \) is
accurate within the resolution limit of the measurement. If \( \gamma < -1 \), the effect of phase distortion on \( \arg F_m(u) \) is less than the effect of truncation distortion and/or measurement noise on \( |F_m(u)| \), so that \( F_m(u) \) is too inaccurate to benefit from the error removal techniques described in section 6.3.2. If \( \gamma > 1 \) then the inequality (6.5) holds.

**Example 1.** \( r_m(x) \) is shown in Fig. 6.5a. The estimate of \( \beta_r \) is 4.1 min. \( f_m(x) \) is shown in Fig. 6.2a (the significance of the negative parts of \( f_m(x) \) is discussed in section 6.2.4). The estimate of \( \beta_m \) is 14 min. The resolution of the measurement is roughly 1.2 min. since the maximum baseline is 1400 wavelengths. Consequently, \( \gamma > 1 \) and so the inequality (6.5) holds. The test has correctly indicated the presence of phase errors in this case since \( N(u) = 0 \) and \( \Psi(u) \) is given by (6.8).

**Example 2.** \( r_m(x) \) is shown in Fig. 6.5b. The estimate of \( \beta_r \) is 3 min. \( f_m(x) \) is shown in Fig. 6.3b. The estimate of \( \beta_m \) is 7.5 min. The resolution of the measurement is roughly 1.0 min. since the maximum baseline is 1650 wavelengths. Consequently, \( \gamma > 1 \) and so the inequality (6.5) holds and the presence of phase errors is predicted. Twiss et al. (1960) do not give estimates of the expected modulus and phase errors in \( F_m(u) \) but it should be noted that \( \arg F_m(u) \) was not measured for \( u > 1250 \) wavelengths. The value of \( \arg F_m(u) \) in the range \( 1650 > u > 1250 \) used to calculate \( f_m(x) \) was obtained by smoothly extrapolating the phase from \( u < 1250 \) and is very likely to contain errors. Further, as mentioned in section 6.1.2, phase measurement with a three element interferometer
(Jennison, 1958) is error sensitive at the longer baselines of a measurement, so the presence of phase errors is quite possible. It is worth noting that \( \beta \) agrees with measurements (Branson, 1965) more recent than those of Twiss et al. (1960).

**Example 2.** \( r_m(x) \) is shown in Fig. 6.5c. The estimate of \( \beta \) is 3.05 min. \( f_m(x) \) is shown in Fig. 3b. The estimate of \( \beta_m \) is 3.5 min. The resolution of the measurement is roughly 0.6 min since \( u_{\text{max}} \approx 2700 \) wavelengths. Consequently, \( |y| < 1 \) so that it is not expected that inequality (6.5) holds and the measurement is accurate within the resolution limit. Fomalont's (1967) estimates of \( \Psi(u) \) and \( |N(u)| \) are shown as error bars in Fig. 6.4b. Using these estimates of the error, the ratio \( \Psi(u) : |N(u)/F(u)| \) is found to be approximately 1.25 : 1, so that phase errors are of the same order as the modulus errors.

### 6.2.4 Negative Parts in the Brightness Temperature

\( f(x) \) is real and non-negative so that \( \hat{f}_m(u) \) must contain significant errors if \( f_m(x) < -V_f \), \( x \) of significant extent. A simple computational test, the theory of which is given in Appendix 3, is presented for establishing whether the inequality (6.5) holds. The computational procedure eliminates those negative parts of the brightness temperature which are due to truncation distortion.

Introduce Fejer's triangular weighting function (Papoulis, 1966), denoted here by \( \text{tri}(u) \), which is defined in the glossary. Define \( \hat{f}_m(x) \) by \( \hat{F}_m(u) \rightarrow \hat{f}_m(x) \), where

\[
\hat{F}_m(u) = \hat{F}_m(u) \text{tri}(u/2u_{\text{max}}) \tag{6.9}
\]

As shown in Appendix 3, \( \hat{f}_m(x) \) can only have negative parts if
N(u) and/or \( \psi(u) \) are appreciable. Define \( \hat{R}_m(u) \) and \( \hat{r}_m(x) \) by

\[
\hat{R}_m(u) = R_m(u) \operatorname{tr}(u/2u_{\max}) \rightarrow \hat{r}_m(x),
\]

which can only have negative parts if \( N(u) \) is appreciable. Notice that the presence of negativeness in \( \hat{r}_m(x) \) is a sufficient, but not necessary, condition for \( N(u) \) to be appreciable.

**Procedure L**

(L.1) Examine \( f_m(x) \), as computed in step (K.2). If \( f_m(x) < -V_f \), \( x \) of significant extent, then the measurement distortion is significant and it is profitable to continue to the next step.

(L.2) Compute \( \hat{r}_m(x) \), using \( \hat{R}_m(u) \rightarrow \hat{r}_m(x) \). If \( \hat{r}_m(x) < -V_r \), \( x \) of significant extent, then \( N(u) \) and/or \( \psi(u) \) must be appreciable. Proceed to the next step to determine whether the inequality (6.5) holds.

(L.3) Compute \( \hat{r}_m(x) \), using \( \hat{R}_m(u) \rightarrow \hat{r}_m(x) \). If \( \hat{r}_m(x) \) is essentially non-negative then \( N(u) \) must be assumed small and, because of step (L.2), \( \psi(u) \) must be appreciable which means that the inequality (6.5) holds and that \( F_m(u) \) can be improved by the processing technique (Procedure M) presented in section 6.3.2. If \( \hat{r}_m(x) < -V_r \), \( x \) of significant extent, then \( F_m(u) \) is too inaccurate to benefit from Procedure M.

**Example 1.** As shown in Fig. 6.6a, \( f_m(x) < -V_T \), \( x \) of significant extent. As shown in Fig. 6.5a, \( r_m(x) \) is essentially positive. So, the inequality (6.5) holds.

**Example 2.** As shown in Fig. 6.6b, \( f_m(x) < -V_T \), \( x \) of significant extent. As shown in Fig. 6.5b, \( r_m(x) \) is essentially positive. So the inequality (6.5) holds.
Example 3. As shown in Figs 6.5c and 6.6c, both \( \hat{f}_m(x) \) and \( \hat{r}_m(x) \) are essentially positive. So, both \( N(u) \) and \( \Psi(u) \) must be assumed small.

6.3 Removal of Phase Errors in Interferograms

If experimental estimates of \( \Psi(u) \) and \( N(u) \) indicate that (or if Procedure K and/or Procedure L has established that) the inequality (6.5) holds then \( f_m(x) \) is not the best available estimate of \( f(x) \). Notice that it is not sufficient merely to neglect any negative parts of \( f_m(x) \) or those parts of \( f_m(x) \) which are of extent greater than \( 2\beta \), (see step (K.1)). Phase errors can alter the whole distribution, as Fig. 6.2a shows. It is now shown how to remove the phase errors from a measured interferogram.

6.3.1 Complex Zeros of the Interferogram

The reconstruction procedure is based on the complex zero theory outlined in sections 5.1.2 and 5.1.3. In these sections it is shown that, given only \( |F(u)| \), all the possible \( f(x) \) can be obtained but the identity of the correct \( f(x) \) is ambiguous. We now show how to compute, using \( |F_m(u)| \) for accurate calculations, with arg \( F_m(u) \) being employed only to resolve this ambiguity, an estimate \( f_r(x) \) which is an improvement on \( f_m(x) \).

The noise threshold for a particular measurement limits the accuracy with which \( f(x) \) can be reconstructed. Bates (1969a) has pointed out that this is equivalent to limiting the number of terms in a trigonometrical Fourier series representation of \( f(x) \). Thus, in (5.5), \( M \) is determined by the number of samples \( F_n \) defined by (5.7) that are above the noise level of \( F_m(u) \). For a particular measurement a suitable
estimate of \( M \) is therefore the maximum integer for which the conditions

\[
\frac{M}{2\beta} \leq u_{\text{max}} \quad (6.10)
\]

\[
\left| F_m \left( \frac{M}{2\beta} \right) \right| > N \left( \frac{M}{2\beta} \right) \quad (6.11)
\]

hold, where \( N(u) \) is the estimated noise level in \( F_m(u) \).

The theory in sections 5.1, 5.2 and 5.3 has been given for the general case in which the source distribution \( f(x) \) is complex. In astronomical situations \( f(x) \) is necessarily real so that Theorem 5 of section 5.3 holds. Then the zeros of \( F(w) \) are situated symmetrically about the imaginary axis and (5.8) and (5.11) show that

\[
P(w) = c \prod_{k=1}^{M} (w-w_k)(w+w_k^*) \quad (6.12)
\]

where

\[
c = F(0) \prod_{m=1}^{M} \left( \frac{m}{2\beta |w_m|} \right)^2 \quad (6.13)
\]

The set \( \{w_k = u_k + jv_k \mid u_k > 0; \quad k = 1,2,\ldots,M \} \) is called the set of complex right-half-plane zeros of \( F(w) \). In special cases some of the \( w_k \) are real, so that as in section 5.1.3, it is convenient to define \( L \) as the number of members of \( \{w_k \mid v_k \neq 0; \quad k = 1,2,\ldots,L < M \} \).

As shown in section 5.3.1, if \( |F(u)| \) or, equivalently, \( R(u) = |F(u)|^2 \) is known accurately, \( 2^L \) distinct possible \( f(x) \) can be defined. It is only necessary to know \( \arg F(u) \) sufficiently accurately to be able to choose the correct \( f(x) \). The set \( \{\zeta_k = p_k + jq_k \mid p_k \neq 0; \quad k = 1,2,\ldots,M \} \) are the complex first quadrant zeros of \( R(w) \). As mentioned in section 5.1.3, in the special cases in which some members of \( \{w_k \} \) are real
there are \( L \) members of \( \{ \zeta_k | q_k \neq 0; k = 1,2, \ldots, L \leq M \} \). If \( \{w_k\} \) and \( \{\zeta_k\} \) are computed accurately from precise versions of \( F(u) \) and \( R(u) \) respectively, then

\[
\begin{align*}
    u_k &= p_k; \quad v_k = q_k \text{ sgn}(v_k) \\
    \end{align*}
\]

When referring to computations performed on data corrupted by measurement distortion, extra subscripts \( m \) and/or \( r \) are appended to \( w_k, u_k, v_k, \zeta_k, p_k \) and \( q_k \) in accordance with the notation introduced in section 6.2.1.

We now describe a procedure (Procedure M) for computing improved brightness temperature distributions from interferograms that contain phase errors. The procedure is based on the theory given in sections 5.1.2 and 5.1.3 and is an extension of the procedure proposed by Bates (1969\textsuperscript{a}).

The procedure is illustrated by the results obtained for the three Examples introduced in section 6.2.2. The numerical analysis techniques used for operating on the examples are described in Chapter 8. A discussion of the theoretical basis of the procedure is given in Appendix 4.

6.3.2 Interferogram Phase Correction Procedure

Procedure M

\( (M.1) \) Obtain \( \beta_R \), an estimate of the halfwidth of \( f(x) \), using step (K.1).

\( (M.2) \) Calculate \( M \), the useful number of Fourier samples, spaced at intervals of \( \frac{1}{2\beta_R} \), available in the measured range of \( F_m(u) \). \( M \) is the largest integer satisfying (6.10) and (6.11) with \( \beta \) replaced by \( \beta_R \).

\( (M.3) \) Starting with \( R_m(u) \), compute the values of the \( M \) members of \( \{\zeta_{mk}\} \). The computational procedure for
computing these zeros is described in section 8.2.1.2.

\[(M.4)\] Compute \( A_m(u) = \left| F_m(u) - P(u)Q(u) \right| \) where \( Q(u) \) is defined in (5.9) and \( P(u) \) is defined in (6.12) with \( \beta \) and \( w_k \) replaced by \( \beta_r \) and \( \zeta_{mk} \) respectively. \( A_m(u) \) is the difference between the actual measured modulus of the interferogram and the modulus of the interferogram defined by the zeros \( \{ \zeta_{mk} \} \). If \( A(u) > N(u) \), \( u \) of significant extent, then it must be assumed that the value of \( M \) obtained in step \((M.2)\) is too small. The resolution limit of the measurement prevents a larger value of \( M \) being used, but we have realized that the closer a zero is to the origin of the complex plane, the more significant is its effect on \( f(x) \). So, it is worth attempting to modify the set \( \{ \zeta_{mk} \} \), as is done in the following step, with \( M \) having the value obtained in step \((M.2)\).

\[(M.5)\] Minimize \( \int_0^{\tilde{u}_{\text{max}}} A^2_m(u) \, du \) by perturbing the values of the members of \( \{ \zeta_{mk} \} \), the final version of which is \( \{ \zeta_{rk} \} \).

The computational procedure for this is described in section 8.2.1.4. Denote the minimized version of \( A_m(u) \) by \( A_r(u) \).

\[(M.6)\] Calculate the position (and denote its value by \( x_0 \)) of the centre of the effective extent of \( f_m(x) \), obtained from step \((K.2)\). If \( |x_0| \) is appreciable it means that \( \arg F_m(u) \) contains a linear term of appreciable magnitude. The expressions in section 5.1.2 have been derived for \( f(x) \) centred at the origin and this linear phase shift, if not removed from \( F_m(u) \), will cause errors in the zeros calculated from the polynomial defined in \((5.10)\). Accordingly, the linear
phase shift is removed in the following step.

(M.7) Compute \( F_{mk}(u) = F_m(u) \exp(-j2\pi x_0u) \).

(M.8) Obtain \( \beta_m \), the halfwidth of \( f_m(x) \), using step (K.2).

Calculate \( K \), the useful number of Fourier samples, spaced at intervals of \( \frac{1}{2\beta_m} \), available in the measured range of \( F_m(u) \). \( K \) is the largest integer satisfying (6.10) and (6.11) with \( M \) and \( \beta \) replaced by \( K \) and \( \beta_m \) respectively. Note that, if it has already been established by Procedure K or Procedure L that the inequality (6.5) holds, then \( \beta_m > \beta \), so that \( K > M \).

(M.9) Starting with \( F_{mr}(u) \), compute the \( K \) members of \( \{w_{mk}\} \), which are the zeros of the polynomial \( P(w) \) defined by (5.10) and (5.7) with \( M \), \( \beta \), and \( F(u) \) replaced by \( K \), \( \beta_m \) and \( F_{mr}(u) \) respectively. The computational procedure required for this step is described in section 8.2.1.1.

(M.10) Compute \( B_m(u) = |F_m(u) - P(u)Q(u)| \) where \( Q(u) \) is defined in (5.9) and \( P(u) \) is defined in (6.12) with \( \beta \), \( M \) and \( w_m \) replaced by \( \beta_m \), \( K \) and \( w_{mk} \) respectively. If \( B_m(u) > N(u) \), \( u \) of significant extent, then minimize \( \int_0^{\text{max}} B_m(u) \, du \) by perturbing the values of the members of \( \{w_{mk}\} \), the final version of which is called \( \{w_{mrk}\} \). Denote the minimized version of \( B_m(u) \) by \( B_r(u) \).

(M.11) As explained in Appendix 4, we must now decide for all \( j \) in the range \( 0 < j < M \) whether \( \zeta_{rj} \) or \( \zeta^*_{rj} \) is a zero of \( F(u) \). We choose \( \zeta_{rj} \) or \( \zeta^*_{rj} \), depending on which is closest to a member of \( \{w_{mrk}\} \). This step is the crux of the procedure. It is based on the two assumptions that \( p_{rk} \) and \( q_{rk} \) are the best available estimates of \( u_k \) and \( |v_k| \) respectively and that the members of the subset \( \{w_{mrk}\} \) defined below are sufficiently close to the
members of \( \{w_R\} \) for one to be able to choose with confidence between \( \zeta_{rk} \) and \( \zeta^*_{rk} \). Fig. 6.7 shows a simple case in which \( M = 1 \) and \( K = 3 \). In this case we would say that \( \zeta_{r1} \) is a zero of \( F(u) \) because it is closest to a member of \( \{w_{mrk}\} \). The selection process can be formalised for any \( M \) and \( K \) as follows.

For a particular integer \( j \), calculate the \( K \) quantities \( C_{jk} = |\zeta_{rj} - w_{mrk}| \) and the \( K \) quantities \( D_{jk} = |\zeta^*_{rk} - w_{mrk}| \). Note the values (call them \( c \) and \( d \)) of \( k \) for which \( C_{jk} \) and \( D_{jk} \) respectively are smallest. Identify an \( M \) member subset \( \{w'_{mrj}\} \) of \( \{w_{mrk}\} \) where \( w'_{mrj} = w_{mrc} \) if \( C_{jc} < D_{jd} \) and \( w'_{mrj} = w_{mrd} \) if \( D_{jd} < C_{jc} \) (no member of \( \{w_{mrk}\} \) may occur more than once in \( \{w'_{mrj}\} \)). The best available estimate of \( \{w_k\} \) is the \( M \) member set \( \{w_{rk} = u_{rk} + jv_{rk}, \ k < M \} \) defined by

\[
\begin{align*}
    u_{rk} &= p_{rk}; \\
    v_{rk} &= q_{rk} \text{ sgn}(v'_{mrk})
\end{align*}
\]

(M.12) Compute \( F_T(u) \) using (5.8), (5.9), (6.12) and (6.13) with the \( w_n \) replaced by \( w_{rn} \) and \( \beta \) replaced by \( \beta_T \).

Compute \( f_T(x) \), the best estimate of \( f(x) \), using (5.5) and (5.7) with \( F(u) \) and \( \beta \) replaced by \( F_T(u) \) and \( \beta_T \).

(M.13) Shift the origin of \( f_T(x) \) to \( x_0 \), determined in step (M.6).

**Example 1.** The positions of computed zeros are shown in Fig. 6.8a. The exact positions of the zeros, \( \{w_k\} \), were computed by the direct search routine described in section 8.2.2, using (6.6) with \( u \) replaced by \( w \). Step (M.2) gives \( M = 3 \). The 3 members of \( \{\zeta_{rj}\} \) computed in step (M.5) are shown, as are the 3 members of \( \{\zeta^*_{rk}\} \). Step (M.8) gives \( K = 11 \), \( x_0 = 0 \). Only 10 members of \( \{w_{mrk}\} \), computed in step (M.11), are shown.
because one member was a great distance off the real axis. The three members of \( \{w_{mk}\} \), which are identifiable by inspection, are marked with double headed arrows. The three members of \( \{w_{rk}\} \) are marked with single-headed arrows. Fig. 6.10a shows \( f_r(x) \) which is seen to be free of negative parts and, except for a small error in the position of the origin, very close to \( f(x) \). arg \( F_r(u) \), shown in Fig. 6.9a, differs from arg \( F(u) \), shown in Fig. 6.2a, only by a linear phase shift term due to the error in the origin of \( F_r(x) \).

Example 2. The positions of computed sets of zeros are shown in Fig. 6.8b. \( M = 3 \) and \( K = 7 \). \( x_0 = 0 \). Fig. 6.9b shows arg \( F_r(u) \), which differs appreciably from arg \( F_m(u) \), shown in Fig. 6.3b, as is expected since \( \Psi(u) \) was estimated to be appreciable by both Procedure K and Procedure L. Fig. 6.10b shows \( f_r(x) \), which is free of negative parts and its extent agrees well with that measured more recently (Branson, 1965).

Example 3. The positions of computed sets of zeros are shown in Fig. 6.8c. Both \( M \) and \( K \) are equal to 5 because \( \beta_m = \beta_r \), as established by Procedure K. \( x_0 = -0.75 \) min. Fig. 6.9c shows arg \( F_r(u) \) which is close to arg \( F_m(u) \) shown in Fig. 6.4b, as is expected since \( \Psi(u) \) was estimated to be small by both Procedure K and Procedure L. Fig. 6.10c shows \( f_r(x) \), which is effectively free of negative parts and agrees reasonably well with Fomalont's (1968) original reconstruction. The structure introduced on the left may be of interest.

6.3.3 Discussion of Phase Correction Procedure

Comparison of Figs 6.2a, 6.3a and 6.4a with Figs 6.10a, 6.10b, 6.10c demonstrates that the procedures presented here can improve the phases of interferograms. In all three cases,
\( f_m(x) \) has negative parts whereas \( f_r(x) \) effectively has none. Also \( f_r(x) \) is narrower in all three cases. There is gratifying agreement between our results for Example 2 and the narrow beam measurements of Branson (1965). As both Procedure K and Procedure L show, the measured phase in Example 3 is accurate within the bounds set by the resolution limit of the measurement, yet even in this case Procedure M gives a slight improvement in that \( f_r(x) \) has less negativeness than \( f_m(x) \).

Figs 6.8a and 6.8b show that Procedure M can be successful even when \( K \) is a few times larger than \( M \). Now, \( K/M \) is the increase in the extent of \( f_m(x) \) over the extent of \( f(x) \), caused by the phase error \( \Psi(u) \). So, it has been demonstrated that severe phase distortion can be compensated for.

The results obtained in Sections 6.2.3 and 6.2.4, by applying Procedure K and Procedure L to the three examples, demonstrate that these procedures afford simple means of checking the quality of measured interferograms.

6.4 The Reconstruction of Brightness Temperatures From Intensity Only Interferograms

6.4.1 Intensity Only Interferograms

As discussed in sections 6.1.2 and 6.1.3 there are cases in both radio and optical interferometry when the phase of the interferogram cannot be measured and only the modulus, \( |F_m(u)| \), or the intensity \( I_m(u) = |F_m(u)|^2 \), of the interferogram is available for reconstruction of the brightness temperature. It is shown in section 5.1.3 that in this case there are in general many different possible brightness temperature distributions. The theory of section 5.1.3 is
now used to obtain all the possible brightness temperature distributions that could have given rise to a particular measured interferogram modulus. The procedure used (Procedure N) is based on the procedure proposed by Bates (1969a). The results of processing three intensity interferograms taken from the literature are used to illustrate Procedure N.

6.4.2 Three Examples

Example 4. This is an interferometer measurement of radio-source P0819-30, made at 468 MHz and 1403 MHz at a position angle of 90°. The measured data and error estimates, obtained by Ekers (1968), are shown in Fig. 6.11. It is assumed that the one dimensional brightness temperature distribution of the source is the same at both frequencies. The interferogram modulus, \(|F_m(u)|\), used for the reconstruction, is shown in Fig. 6.11 as the full line drawn smoothly through the measured points. The measured points are sufficiently close together to be able to choose \(|F_m(u)|\) with confidence.

Examples 5 and 6. These examples show the application of the reconstruction technique to optical intensity interferograms obtained with the stellar interferometer at Narrabri (Hanbury-Brown et al., 1967a and b). The interferometer at Narrabri has so far been used only to measure the angular diameters of single stars (Hanbury-Brown et al., 1967b) and the diameters and separations of the stars in a multiple star system (Hanbury-Brown et al., 1970). The possibility of obtaining information about the actual structure of the brightness temperature distribution to indicate, for example, limb darkening or rotational distortion, has been examined (Hanbury-Brown et al., 1967a and 1967b, Johnson and Wareing,
1970). Measurements of these sorts of phenomena would be of considerable value for verifying proposed theoretical stellar models. It appears that, in general, the accuracy of the Narrabri instrument is at present not sufficient to determine anything other than the diameter of a circular disc of constant brightness temperature assumed equivalent to the star being observed. However, because very severe rotational distortion could possibly be detected by the present Narrabri instrument (Johnson and Wareing, 1970), and in anticipation of a more accurate instrument being constructed, it is considered worthwhile to demonstrate the application of the reconstruction technique to data obtained at Narrabri.

Example 5 is a measurement of the intensity of the interferogram of αPsA measured at a wavelength of 4608Å and Example 6 is a measurement of the intensity of the interferogram of βCru measured at a wavelength of 4385Å. The measured data and the estimated experimental errors, obtained by Hanbury-Brown et al. (1967b), are shown in Fig. 6.12. The interferogram intensities, |F_m(u)|, used in the reconstruction procedure were assumed to be the smooth curves passing through the mean values of the measured data points shown in Fig. 6.12.

6.4.3 Intensity Interferogram Reconstruction Procedure

Procedure N - Reconstruction of all possible brightness temperatures defined by an intensity interferogram.

(N.1) Compute r_m(x) and β_m using step (K.1) (section 6.2.3).

(N.2) Compute f_m(x) using step (L.3) (section 6.2.4). As explained in Appendix 3, f_m(x) cannot have negativeness unless N(u), the error in |F_m(u)|, is significant.
Thus, if \( F_m(x) \) is negative \( N(u) \) is significant and little confidence can be placed in the brightness temperature reconstructions. If \( F_m(x) \) is essentially positive, it is worthwhile continuing with the reconstruction procedure.

(N.3) Calculate \( M \), using step (M.2).

(N.4) Calculate the \( M \) members of \( \{ \varsigma_{mk} \} \) using step (M.3).

(N.5) Using step (M.4), refine the positions of the zeros \( \{ \varsigma_{mk} \} \) to give the more accurate set \( \{ \varsigma_{rk} \} \).

(N.6) Compute \( F_r^{(0)}(u) \), the interferogram (modulus and phase) corresponding to \( R_m(u) \) that has all its zeros in the upper half plane. \( F_r^{(0)}(u) \) is computed from (5.8), (5.9) (6.12) and (6.13) with \( w_n \) and \( \beta \) replaced by \( \varsigma_{rn} \) and \( \beta_r \) respectively.

(N.7) If \( L \) of the \( M \) members of \( \{ \varsigma_{rk} \} \) are complex, \( L < M \), set up the \( 2^L \) sequences \( \{ \tau_p \} \) defined in section 5.1.3.

(N.8) The zero flipping procedure described in 5.1.3 is now used to compute all the possible interferograms \( F_r^{(m)}(u) \), that have the same modulus \( |F_m(u)| \). Since \( f(x) \) is real, \( F(u) = F^*(u) \) and it follows from (5.8) and (6.12) that if a zero \( w_n \) is flipped, the zero \(-w_n^*\) must also be flipped. Compute the possible \( F_r^{(m)}(u) \) using

\[
F_r^{(m)}(u) = F_r^{(0)}(u) \prod_{p=1}^{L} \frac{u - \tau_p \varsigma_{p}^*}{u - \tau_p \varsigma_{p}^{**}} \frac{u + \tau_p \varsigma_{p}^*}{u + \tau_p \varsigma_{p}^{**}} \tag{6.15}
\]

where the \( \tau_p \) in (6.15) are taken from the \( m^{th} \) of the \( 2^L \) sequences \( \{ \tau_p \} \). The sequence \( \{ \tau_p \} \) is defined in section 5.1.3. Note that it is only necessary to compute the \( F_r^{(m)}(u) \) at the sampling points required in
the next step.

(N.9) Construct the possible brightness temperature distributions \( f_r^{(m)}(x) \) using (5.5) and (5.7) with \( F(u) \) and \( \beta \) replaced by \( F_r^{(m)}(u) \) and \( \beta_r \). Notice that only \( 2^{L-1} \) instead of \( 2^L \) different brightness distributions need be calculated since \( f_r^{(m)}(-x) \), the mirror image of \( f_r^{(m)}(x) \), can be obtained by exchanging all the zeros of \( F_r^{(m)}(w) \) for their complex conjugates.

(N.10) Discard as unphysical all those \( f_r^{(m)}(x) \) which have negative values. The remaining \( f_r^{(m)}(x) \) and their mirror images are all the possible brightness temperature distributions that have width \( 2\beta_r \) and interferogram modulus \(|F_m(u)|\).

**Example 4.** \( r_m(x) \) and \( r_m^*(x) \) are shown in Fig. 6.13. Notice that \( r(x) \) is wholly positive so that the slight negativeness in \( r_m(x) \) can be attributed to truncation of \(|F_m(u)|\). The estimate of \( 2\beta_r \) obtained from inspection of \( r_m(x) \) is 9.4 min. With the maximum baseline \( u_{\text{max}} = 2200 \) wavelengths, this value of \( \beta_r \) gives \( M = 6 \). The six members of \( \{\zeta_{r_k}\} \) are shown in Fig. 6.14 above the real axis; \( \{\zeta_{r_k}^*\} \) are below the real axis. Since only 5 members of \( \{\zeta_{r_k}\} \) are complex, there are 32 different possible \( f_r^{(m)}(x) \) that could have given rise to \(|F_m(u)|\) shown in Fig. 6.11. Only 16 of these need to be computed, the remaining 16 being mirror images of them. Fig. 6.15 shows 10 of the possible distributions, the remaining six being similar to some of the ten shown. The particular combinations of zeros giving rise to each distribution is included in Fig. 6.15 by indicating the signs of the imaginary parts of each zero. Thus, + - + + - indicates that
the first, third and fourth zeros, ordered according to distance from the imaginary axis, are above the real axis, and the second and fifth zeros are below the real axis.

Of the ten distributions shown in Fig. 6.15, only the three shown in Figs 6.15a, 6.15d and 6.15h have negligible negative parts. The rest may be discarded as physically impossible. There is no way, without the phase of the interferogram, of determining which of the distributions of Figs 6.15a, 6.15d and 6.15h, or their mirror images, is the correct distribution. In fact, in this case, the correct distribution is that of Fig. 6.15h. There is close agreement of the size and structure of this distribution with that of Fomalont (1968) which was obtained using both the measured modulus and phase of the interferogram. Notice that although the original estimate of $2\beta_\tau = 9.4$ min. arc. was somewhat greater than the more accurate value of 6 min. arc. obtained by Fomalont (1968), the final reconstruction is essentially zero outside $2\beta = 6$ min. arc., showing that in this case a choice of $2\beta$ too large has not prevented a good reconstruction. Although the distribution in Fig. 6.15a is not greatly different from the correct distribution, the fact that the distribution of Fig. 6.15d is quite different emphasizes that there is a real danger in practice of obtaining incorrect reconstructions from modulus only measurements.

Examples 5 and 6. $r_m(x)$ and $f_m(x)$ for these two examples are shown in Fig. 6.16. For both cases, the fact that $f_m(x)$ is wholly positive indicates that the negativeness in $r_m(x)$ is due mainly to truncation of the interferogram so that it is worthwhile continuing with the reconstruction procedure.
Examination of the $r_m(x)$ in Fig. 6.16 gives estimates for $2\beta_r$ of $4 \times 10^{-3}$ secs.arc. for $\alpha$ PsA and $2.5 \times 10^{-3}$ secs.arc. for $\beta$ Cru. $M = 2$ for $\alpha$ PsA and $M = 4$ for $\beta$ Cru. The members of \{\zeta_{rk}\} are shown in Fig. 6.17 above the real axis. \{\zeta^*_{rk}\} are below the real axis. Since three of the members of \{\zeta_{rk}\} for $\beta$ Cru are complex, there are eight possible brightness temperature distributions. Four of these are shown in Fig. 6.18, the remaining four are their mirror images. One of the zeros of $\alpha$ PsA has a very small imaginary part so that moving it across the real axis has negligible effect on $f_r^{(m)}(x)$. Thus for $\alpha$ PsA, only the single distribution shown in Fig. 6.19 and its mirror image are possible.

None of the $f_r^{(m)}(x)$ calculated for $\alpha$ PsA or $\beta$ Cru are wholly positive. Procedure L can be used to determine whether this negativeness is due to the presence of phase errors in $F_r^{(m)}(u)$ or only to the truncation of $F_r^{(m)}(u)$. Thus, when $f_r^{(m)}(x)$ is calculated using step (N.9) with $F_r^{(m)}(u)$ replaced by $\hat{f}_r^{(m)}(u) = \text{tri}(u/2u_{\text{max}})F_r^{(m)}(u)$, the distribution shown in Fig. 6.18a for $\beta$ Cru and in Fig. 6.19 for $\alpha$ PsA give $f_r^{(m)}(x)$ that are wholly positive indicating that the negativeness in these distributions is due only to the truncation of $F_r^{(m)}(u)$. The remaining distributions shown in Fig. 6.18b, 6.18c, 6.18d for $\beta$ Cru all give negative $f_r^{(m)}(x)$. Thus, for $\beta$ Cru, the distribution shown in Fig. 6.18a can confidently be chosen as the correct one for the $|F_m(u)|$ shown in Fig. 6.12.

Notice that the widths of the significant parts of the distributions in Fig. 6.18a and Fig. 6.19 are in good agreement with the stellar diameters originally obtained by Hanbury-Brown et al. (1967b), which are shown for comparison.
The fact that the reconstructed brightness temperature distributions are not zero outside the stellar widths can be attributed to the uncertainty in the measured data and the limited range of available baselines.

6.4.4 Discussion of Intensity Only Reconstruction

An important feature of the reconstruction procedure described in section 6.4.3 is that it is not a model fitting technique. The reconstruction methods originally used by those who obtained the data used in the Examples (Ekers, 1968; Hanbury Brown et al., 1967a and b) were model fitting techniques. Ekers assumed that the brightness temperature distribution under observation was composed of two gaussian distributions, whilst in the optical case Hanbury-Brown et al. assumed that the stars were uniform discs with specified limb darkening. Structure present in the observed sources but not consistent with these models could obviously not be included in the reconstructed brightness temperature distributions. The only assumption about the source that is required with Procedure N is that it is effectively of finite extent, an assumption valid in most cases as is discussed in section 6.1.1.

Both procedure M and procedure N require that $|F_m(u)|$ be known for values of $u$ spaced $1/4\beta$ apart. Although interferograms are sometimes measured as continuous functions of $u$ (Ekers, 1968), it is more usual for the interferogram to be sampled at discrete values of $u$. In such cases, as in the Examples described above, the interferograms can be obtained by steering smooth curves through the measured points. In cases for which, because of coarse baseline sampling or large
uncertainty in the measurement, alternative interpolations of equal plausibility are possible, it would be necessary to carry out the reconstruction procedure for different possibilities to determine the likely spread of the features of the reconstructed distribution (Maltby and Moffet, 1962).

Using model fitting reconstruction techniques of the type described by Ekers (1968), a considerable amount of computational effort would be required to compute all the possible brightness temperature distributions corresponding to a particular measured interferogram modulus, and there would be no way of telling when all possible distributions had been found. Procedure N allows all possible distributions to be obtained directly. In this respect the present method could provide a useful complement to modelling methods. When a best fit interferogram modulus has been found with a modelling method, Procedure N could be used to determine all the other possible brightness distributions having the same interferogram modulus. Unless some extra information is available, there would be no basis for choosing one of these distributions before any of the others.
FIGURE 6.1a TWO ELEMENT INTERFEROMETER

FIGURE 6.1b MODULUS AND PHASE ERRORS IN $F(u)$. 
FIGURE 6.2a. BRIGHTNESS TEMPERATURE DISTRIBUTION FOR EXAMPLE 1.

\[ - F(x) \quad \text{---} \quad F_m(x) \]

FIGURE 6.2b. INTERFEROGRAM FOR EXAMPLE 1.

\[ |F(u)| \quad \text{---} \quad \arg F(u) \quad \text{---} \quad \arg F_m(u) \]
Figure 6.3a. Brightness temperature distribution for example 2.

Figure 6.3b. Interferogram for example 2.

--- $|F_m(u)|$  --- $\arg F_m(u)$
Figure 6.4a. Brightness Temperature Distribution for Example 3.

Figure 6.4b. Interferogram for Example 3.

--- $|F_m(u)|$  × Fomalont  ○ Ekers
--- $\arg F_m(u)$  □ Fomalont
FIGURE 6.5. BRIGHTNESS TEMPERATURE AUTOCORRELATION FUNCTIONS FOR STEPS (K.1) AND (L.3) — $r_m(x)$, $r_m(x)$
(a) EXAMPLE 1. (b) EXAMPLE 2. (c) EXAMPLE 3.
Figure 6.6. Brightness Temperature Distributions $\hat{f}_m(x)$ for Step (L.2)
(a) Example 1. (b) Example 2. (c) Example 3.
Figure G.7 Separation between the members of 
\[ \{ \zeta_{rN} \} \cup \{ \zeta^{*}_{rK} \} \] and the members of \[ \{ w_{mrk} \} \]
FIGURE 6.8a.

FIGURE 6.8b.

FIGURE 6.8c.

FIGURE 6.8. COMPUTED ZEROS OF INTERFEROGRAMS
(a) EXAMPLE 1. (b) EXAMPLE 2. (c) EXAMPLE 3.
Δ EXACT POSITIONS OF ZEROS \{w_k\} (ONLY FOR EXAMPLE 1)
○ \{\delta_{rk}\} ABOVE REAL AXIS, \{\delta_{rk}\} BELOW REAL AXIS, FOR STEP (M.5)
+ \{\text{work}\} COMPUTED IN STEP (M.10.)
≠ \{\text{work}\} CHOSEN IN STEP (M.11.) ≠ \{\text{work}\} CHOSEN IN STEP (M.11.)
FIGURE 6.9. THE PHASE OF THE RECONSTRUCTED INTERFEROGRAMS.
(a) EXAMPLE 1. (b) EXAMPLE 2. (c) EXAMPLE 3.
FIGURE 6.10. RECONSTRUCTED BRIGHTNESS TEMPERATURE
(a) EXAMPLE 1. (b) EXAMPLE 2. (c) EXAMPLE 3.
Figure G.11. Interferogram for example 4, measurement made by Ekers (1968)

\( \frac{\lambda}{468 \text{ MHz}}, \frac{\lambda}{1403 \text{ MHz}} \) - modulus of interferogram used for reconstruction.
FIGURE 6.12. INTENSITY INTERFEROGRAMS MEASURED BY HANDBURY BROWN et al. (1967b) (a) EXAMPLE 5 (b) EXAMPLE 6.
- MEASURED POINTS, ----- $|F(u)|^2$ FOR RECONSTRUCTION.
FIGURE 6.13. BRIGHTNESS TEMPERATURE AUTOCORRELATIONS FOR EXAMPLE 4.

FIGURE 6.15. POSSIBLE BRIGHTNESS TEMPERATURE DISTRIBUTIONS FOR EXAMPLE 4.
FIGURE 6.16a.

FIGURE 6.16b.

FIGURE 6.16  BRIGHTNESS TEMPERATURE AUTOCORRELATION FUNCTIONS
(a) EXAMPLE 5  (b) EXAMPLE 6

--- \( r(x) \), F.T. of \( |F(u)|^2 \); --- \( \hat{r}(x) \), F.T. of \( |F(u)|^2 \) \( \text{erf}(u/2u_{\text{max}}) \)
FIGURE 6.17a.

FIGURE 6.17b.

FIGURE 6.17 COMPLEX ZEROS OF $r(w)$.
(a) EXAMPLE 5  (b) EXAMPLE 6.
Figure G.18. Possible brightness temperature distributions for example G.
FIGURE C.19. BRIGHTNESS TEMPERATURE DISTRIBUTION FOR EXAMPLE 5.
W IS THE STELLAR DIAMETER OBTAINED BY HANBURY-BROWN ET AL. (1967b)
CHAPTER 7. Combination of Holographic and Complex Zero Reconstruction Techniques

In previous chapters two approaches to the problem of obtaining phase information for the reconstruction of radiating source distributions have been discussed. In chapters 2, 3 and 4 a holographic approach was considered. In chapters 5 and 6 application of the theory of entire functions was shown to be useful. In this chapter reconstruction procedures that combine these two approaches are discussed.

7.1 Introduction

The work in Chapters 2, 3 and 4 is based on the holographic principle of observing a source \( f(x) \) that is composed of a known part \( g(x) \) and an unknown part \( h(x) \). The work in Chapters 5 and 6 is based on the fact that the F.T. of any source of finite extent is characterized by the positions of its zeros in the complex plane. In this chapter these two concepts are combined to provide solutions to further problems of source reconstruction. In this chapter, the source distribution \( f(x) \) and its F.T. \( F(u) \) satisfy (2.1)-(2.4) and (2.12)-(2.13). The theory given in sections 5.1.1-5.1.3 also applies to \( f(x) \) and \( F(u) \). Since \( g(x) \) and \( h(x) \) are of finite extent, \( G(w) \) and \( H(w) \), as well as \( F(w) \) are entire functions characterized by their zeros.
7.2 Use of an A Priori Estimate of the Source Distribution

7.2.1 Description of the Problem

The problem examined in this and the next two sections may be stated as follows: given the measured modulus of the F.T. of a source distribution and an a priori estimate of the source distribution, determine the actual source distribution. Situations in which the solution of this problem would be of value occur, for example, in the fields of antenna measurement radio interferometry and radar.

Suppose that a newly manufactured antenna has the modulus of its radiation pattern measured and it is found that there are minor errors in the pattern. As explained in section 7.2.2, knowledge of the aperture distribution of the antenna will be of value in correcting the errors. Provided the errors in the antenna are not too large, a good estimate of the actual aperture distribution will be the distribution that the antenna was designed to have. Then, if a solution to the problem given in the first sentence of this section can be found, the actual aperture distribution of the antenna will be obtainable from only the modulus of the radiation pattern of the antenna.

The brightness temperature distributions of some radio astronomical sources vary with time. In some cases it is important to observe these variations closely (Stubbs, 1971). The solution of the problem mentioned above would allow the distribution known at any time to be used as an initial estimate for the distribution at some later time, allowing the actual brightness temperature distribution to be determined from only the modulus of the interferogram. This
process could only be successful, of course, if the very first brightness temperature distribution is determined accurately using both the modulus and phase of the interferogram and if the time between successive interferogram modulus measurements was small enough so that changes in the brightness temperature distribution were not too large.

Under certain circumstances (cf. Bates 1969c) the silhouette of a radar target is given approximately by the F.T. of the backscattering pattern of the target. If the backscattering pattern of an artificial satellite was measured before the satellite was launched into orbit then a small change in shape, occurring after launch, could be detected by examining an extended sequence of radar returns. Although most satellite tracking radars are of the coherent type and can be used for making accurate phase measurements, nevertheless, as has been pointed out elsewhere (Bates 1969c), it is questionable whether the phase variations of the scattering pattern can be abstracted from the total phase variations of the radar returns. So, only the modulus of the scattering pattern is observable unequivocally. Solution of the problem mentioned above will allow changes in the satellite structure to be reconstructed.

7.2.2 A Reconstruction Procedure Using an A Priori Estimate of a Source Distribution and the Measured Modulus of its F.T.

Consider a source distribution f(x) of finite extent 2\(\beta\). f(x) and its F.T. \(F(u)\) satisfy (5.2) and (5.3). Further, let f(x) be composed of a known part g(x) and an unknown part h(x). Then f(x) and F(u) satisfy (2.10) and (2.11). f(x) and F(u) are not necessarily an antenna aperture distribution.
and its radiation pattern, but can in general be any source distribution and its F.T. provided the latter is a physically observable quantity.

The problem posed at the start of the previous section may now be formulated as follows: given \( g(x) \) and \( |F(u)| \), determine \( f(x) \) when \( h(x) \ll g(x) \), where \( g(x) \) is the a priori estimate of \( f(x) \) and \( h(x) \) is the difference, assumed small, between \( f(x) \) and \( g(x) \). The solution to this problem, proposed by Bates (1971), provides the basis for the reconstruction procedure (Procedure 0) given below. The basis of Bates' (1971) suggestion is that, if \( h(x) \) is small enough, \( F(w) \) will be close to \( G(w) \) and, in particular, the zeros of \( F(w) \) will be close to the zeros of \( G(w) \). \( F(w) \) has zeros \( \{w_n = u_n + jv_n\} \) as described in section 5.1.2. Suppose that \( G(w) \) has zeros \( \{w_{gn} = u_{gn} + jv_{gn}\} \). From a knowledge of \( |F(u)| \), the set of zeros \( \{\zeta_n = p_n + jq_n\} \) of \( R(w) \), where \( R(w) \) is defined in (5.17) can be found, as described in section 5.1.3. Then if \( H(w) \) is to be the smallest possible perturbation of \( G(w) \)

\[
\begin{align*}
  u_n &= p_n \\
  v_n &= q_n \text{sgn}(v_{gn}).
\end{align*}
\]

Since the \( w_n \) are now determined, \( F(w) \) is completely defined (except possibly for a linear phase shift term) and \( f(x) \) can be obtained from (5.3). An actual reconstruction procedure based on this process is now detailed. In general, the measured quantity \( |F(u)| \) will contain the same types of noise and truncation distortion as is contained in measured interferograms, so it is convenient to use the notation described in section 6.2.1 to distinguish between actual, measured and
computed quantities. Some of the steps in Procedures K, L, M and N are also used. It must be remembered, however, that \( f(x) \) and \( F(u) \) are not necessarily a brightness temperature - interferogram pair, i.e. \( f(x) \) is not necessarily real so that the zeros of \( F(w) \) need not be symmetrical about the imaginary axis.

Procedure 0

(0.1) Compute \( r_m(x) \) and \( \beta_r \) using step (K.1) (section 6.2.3).

(0.2) Calculate \( M \) using step (M.2).

(0.3) Starting with \( R_m(u) \), compute the \( 2M \) members of

\[
\{ \xi_{mk} = p_{mk} + jq_{mk} \}.
\]

(0.4) Since \( g(x) \) is known, \( G(w) \) can be obtained using \( g(x) \rightarrow G(w) \). If an analytic expression can be obtained for \( G(w) \), use the simple search routine described in section 8.2.2 to compute the zeros

\[
\{ w_{gk} = u_{gk} + jv_{gk} \}
\]

of \( G(w) \). If an analytic expression for \( G(w) \) cannot be obtained, which will usually be the case, obtain \( 2M \) members of \( \{ w_{gk} \} \) as the zeros of the polynomial \( P(w) \) defined in (5.10) with \( F(u) \) and \( \beta \) replaced by \( G(u) \) and \( \beta_g \), where \( 2\beta_g \) is the extent of \( g(x) \).

(0.5) Obtain the \( 2M \) of the members of \( \{ w_k = u_k + jv_k \} \), the zeros of \( F(w) \), using (7.1) and (7.2) with \( \{ \xi_{mk} \} \) and \( \{ w_{gk} \} \).

(0.6) Compute \( F_r(u) \) using (5.8), (5.9), (5.11) and (5.12), with \( \beta \) replaced by \( \beta_r \). Compute \( f_r(x) \), the best estimate of \( f(x) \), using (5.5) and (5.7) with \( F(u) \) and \( \beta \) replaced by \( F_r(u) \) and \( \beta_r \).

The crux of Procedure 0 is step (0.5). An important question is how much smaller than \( g(x) \) must \( h(x) \) be to ensure
that (7.2) holds. No general answer has been found to this question. Bates (1971) points out that in order that
\( \text{sgn}(v_n) = \text{sgn}(v_{gn}) \), there must exist \( \{ \text{real } \alpha_n > 0, \alpha_n < |v_{gn}| \} \) such that

\[
|G(w)| > |H(w)| \quad \text{on} \quad |w - w_{gn}| = \alpha_n
\]  

(7.3)

This follows from the fact that \( G(w_n) = -H(w_n) \) and from \( G(w) \) and \( H(w) \) being analytic everywhere in the finite \( w \)-plane (Morse and Feshbach, 1953, Pg. 368). (7.3) determines how small \( h(x) \) must be in comparison with \( g(x) \) but is of little use for placing useful limits on \( h(x) \) for a particular \( g(x) \). However, Procedure 0 is self-checking in the sense that if \( h(x) \) is too large, it is unlikely that the choice of members of \( \{w_k\} \) in step (0.5) will be unambiguous. The number of members of \( \{w_k = u_k + jv_k\} \) for which \( \text{sgn}(v_k) \) is ambiguous will determine the confidence that can be placed in \( f(x) \).

Two types of \( g(x) \) will be unsuited for use in Procedure 0. Consider first (Fig. 7.1a) a case in which a zero of \( G(w) \) is situated at \( [u_g + jv_g] \) on, or very near, the real axis. The corresponding zero of \( F(w) \), \( [u_1 + jv_1] \), will also be near the real axis since zeros of \( F(w) \) are close to those of \( G(w) \). However, the zeros of \( R(w) \) which are the ones actually computed, will exist at \( [u_1 + jv_1] \) and \( [u_1 - jv_1] \). The distances between each of these zeros and \( [u_g + jv_g] \) is very nearly the same for both and it will not be possible to choose with certainty the zero of \( F(w) \).

Consider next (Fig. 7.1b) a case in which two zeros of \( G(w) \) are positioned symmetrically about the real axis. Let these zeros be at \( [u_g + jv_g] \) and \( [u_g - jv_g] \). \( F(w) \) will have
zeros close to these, say, \([u_1+jv_1]\) and \([u_2-jv_2]\). The 
zeros of \(R(w)\) will exist at \([u_1+jv_1]\), \([u_1-jv_1]\), \([u_2+jv_2]\), \([u_2-jv_2]\). It will not be clear which of these zeros 
represent the smallest possible perturbation of the two zeros 
of \(G(w)\) and again the zeros of \(F(w)\) cannot be chosen with 
certainty. For Procedure 0 to be useful, therefore, the 
zeros of \(G(w)\) should be appreciably displaced from the real 
axis and should be positioned asymmetrically about it. In 
general, the more structure there is in \(g(x)\), the further 
away from the real-axis are the zeros of \(G(w)\). From theorem 
6 of section 5.5, the zeros of \(G(w)\) will be positioned 
symmetrically about the real axis if \(g(x)\) is symmetric and 
has uniform phase. So we have the rather surprising result that 
the more complicated and asymmetrical a source is, the more 
likely it is that Procedure 0 will give a correct reconstruction.

7.2.3 An Example of the Use of Procedure 0.

Most antennas are required to have a single main beam 
and low sidelobes. To satisfy this requirement the aperture 
distributions are usually designed to be symmetric and have 
uniform phase so that, as explained at the end of the last 
section, Procedure 0 is not suitable for measuring antenna 
aperture distributions directly. However, the test antenna 
and a fixed reference antenna can be combined to provide an 
asymmetrical source distribution. All that is necessary is to 
displace the centre of the test antenna from the axis of the 
turntable by a small fraction of the extent of the test 
antenna. As explained in section 3.4.1, the fixed reference 
antenna effectively provides a point source at the centre of
the turntable. The combination of this point source and the off-axis test antenna gives an asymmetric distribution. A simple, ideal, radiation pattern, D(u), can be assumed for the pattern of the antenna under test. This may be obtained from a knowledge of the ideal design aperture distribution of the test antenna or from inspection of the modulus of its radiation pattern. If the antenna is situated a distance $T_T$ from the centre of the turntable, where $T_T$ is a small fraction of $2\beta$, the width of the antenna, and if the ideal pattern $D(u)$ is combined with a fixed reference signal of strength $S$, then the a priori estimate of the pattern, $G(w)$, is

$$G(w) = S + D(w) \exp(j2\pi T_T w) \quad (7.4)$$

An example is now given of the use of Procedure O. A reference signal is included to give an asymmetrical source distribution. The aperture distribution of a circular electrostatic acoustic transducer having a diameter of 3.9 cm was obtained using the method. The experimental set-up is described in Section 3.1. The 3.9 cm diameter transducer was placed on the turntable in a position such that $T_T = 1.25$ cm. $|F(u)|$, the modulus of the radiation pattern of the reference signal and test signal combined is shown in Fig. 7.2. The modulus of the radiation pattern of the antenna by itself is shown in Fig. 7.4. Assume that the ideal pattern is that due to a source of constant amplitude having an extent necessary to produce a main beam having the same width as that of the test antenna. Then inspection of Fig. 7.4 gives

$$D(u) = \frac{\sin \frac{3\pi u}{3\pi \eta}}{3\pi \eta} \quad (7.5)$$
and the ideal combined pattern is

\[ G(w) = 0.404 + \sin \left( \frac{3\pi w}{3\pi w} \right) \exp(j5\pi w/2) \]  

(7.6)

where the strength of the constant reference signal has been obtained by inspecting \(|F(u)|\) in Fig. 7.2. The zeros of \(G(w)\), obtained using the simple search routine described in section 8.2.2, are shown in Fig. 7.3. The zeros of \(R(w)\), computed using the F.T. procedure described in section 8.2.2, are also shown in Fig. 7.3. The zeros of \(F(w)\) are obviously the zeros of \(R(w)\) below the real axis. It will be noted that no zeros of \(R(w)\) could be found corresponding to the zeros of \(G(w)\) in the region \(v = 2.0\). This was due to computational difficulties which are described in section 8.2.2. For reconstruction purposes, the zeros of \(F(w)\) for \(v \approx 2.0\) were assumed to be identical to those of \(G(w)\): little error would be introduced by this assumption owing to the relatively large displacement of these zeros from the visible region of the pattern, which is \(v = 0, |u| < 1\). The reconstructed pattern, with its phase, is shown in Fig. 7.2 and its aperture distribution \(f(x)\) is shown in Fig. 7.5. The reconstructed pattern with the reference source removed is shown in Fig. 7.4 with the actual measured pattern for comparison. The main features of the pattern are reproduced, although the differences in the outside sidelobes indicate that there are errors in the reconstructed \(f(x)\). This could possibly indicate that the ideal pattern \(D(u)\) assumed in (7.5) was not sufficiently close to the actual pattern, so that the zeros of \(G(w)\) are not close enough to the zeros of \(F(w)\) for step (0.5) to be successful. This appears unlikely, however,
because of the similarity between the zeros of $G(w)$ and $R(w)$ evident in Fig. 7.3. The most likely source of the error is in the positions of the zeros of $R(w)$ outside the range of the measurement. Because $|F(u)|$ was only measured for $|u| < 1$, all zeros of $R(w)$ with real parts greater than 1.0 have effectively been obtained by extrapolation, a process which is well known to be error sensitive.

7.3 **Application of the Holographic Principle to Interferometry**

If single stars or continuous distributions of brightness temperature are observed using an interferogram then, as explained in section 6.1.1, the phase of the interferogram must be known for unambiguous computation of the brightness temperature, except in the special cases mentioned in section 5.1.3. Suppose, however, that two separated distributions of brightness temperature are observed together. Each distribution (or star) can serve as a reference for the other, so that by appealing to the concept of holography it is reasonable to look for a method of unique reconstruction using only the intensity of the interferogram.

Consider a radio or optical astronomical brightness temperature distribution $f(x)$ that is composed of two individual distributions $g(x)$ and $h(x)$, so that (2.2) is satisfied. $F(u)$, $G(u)$ and $H(u)$, the interferograms corresponding to $f(x)$, $g(x)$ and $h(x)$ satisfy (2.4). Then the theory of section 2.3.1 shows that, if $g(x)$ and $h(x)$ are separated by a distance greater than the extent of either of them, then $g(x)\Theta h^*(x)$ can be obtained directly from the F.T. of $|F(u)|^2$. If $g(x)$ is known $h(x)$ can be obtained from $g(x)\Theta h^*(x)$ using the method described in section 2.3.1.
In particular if the reference source, $g(x)$, is a point source the theory of section 2.2.3 shows that $h(x)$ can be obtained directly from the F.T. of $|F(u)|^2$. This condition will hold approximately if the width of $g(x)$ is much less than the width of $h(x)$. This can be guessed from an inspection of the centre part of $r(x)$, if its central part consists of a narrow peak on a wide pedestal. The idea of using a point reference source in interferometry has also been suggested recently by Goodman (1970b) and Kohler and Mandel (1970). It has been demonstrated to be practical in the laboratory by Beard (1969).

When $g(x)$ cannot be considered to be a delta function, some sort of deconvolution procedure is required to obtain $h(x)$ from $g(x) \ast h^*(x)$. In section 3.2.1 it is shown how to take the finite width of $g(x)$ into account, provided the major lobe of $G(u)$ is somewhat wider than the interval containing the significant structure of $H(u)$. In an astronomical situation it may not be possible to choose $g(x)$ much smaller than $h(x)$ and, if any information about $G(u)$ by itself is available at all, it is probable that only $|G(u)|$ will be known. In this case the deconvolution procedure is more difficult than that used in section 3.2.1. In fact a completely new method of deconvolution is required, as is described in detail in the following sections.

7.4 Deconvolution Using Complex Zeros

7.4.1 Statement of the Problem

The problem considered in this and the next three sections can be stated as follows: given $g(x) \ast h^*(x)$, the cross correlation of two functions $g(x)$ and $h(x)$ that are of finite extent, and given $|G(u)|$, the modulus of the F.T. of
one of them, determine $g(x)$ and $h(x)$. Although in this case we start with the correlation function, the problem is essentially the same as starting with the convolution of $g(x)$ and $h(x)$. The problem is therefore a special case of the problem known in the literature as the "deconvolution" problem in which $h(x)$ must be obtained when $g(x) * h(x)$ and $g(x)$ are known. Knowledge of $g(x)$ implies complete knowledge of $G(u)$; for the new technique only $|G(u)|$ is needed.

As in section 2.3.1, it is convenient to define $e(x)$ as the correlation of $g(x)$ and $h(x)$.

$$e(x) = g(x) * h^*(x) \quad (7.7)$$

and

$$E(u) = G(u) H^*(u) \quad (7.8)$$

### 7.4.2 Review of Methods of Deconvolution

The deconvolution problem occurs in many situations. For example, it occurs in such diverse fields as determining the structure of astronomical sources using radio telescopes (Bracewell and Roberts, 1954), determining the frequency response of antennas from measurements made in the time domain (Nicolson, 1968), processing mass spectrometer data (Morrison, 1963) and optical image restoration (Stroke, 1969).

Techniques for carrying out deconvolution can conveniently be grouped according to whether the deconvolution is carried out in the $x$ domain or the $u$ domain.

The simplest form of deconvolution in the $x$ domain is the serial process described by Bracewell (1965, p.35), in which $h(x)$ is progressively calculated starting from one end,
as sampled versions of $e(x)$ and $g(x)$ are moved serially through each other. This technique is of little practical use because any noise present in $e(x)$ or $g(x)$ tends to compound rapidly as the deconvolution proceeds giving an $h(x)$ that is completely wrong. MacAdam (1970) has shown that the method can be useful if suitable constraints can be placed on $e(x)$, $g(x)$ and $h(x)$. Morrison (1963) mentions two $x$-domain deconvolution techniques, one of which is suitable for cases in which $e(x)$ has well defined maxima and the other is useful when $h(x)$ is a slowly varying function. Burr (1955) gives a useful survey of methods for deconvolution in two dimensions.

The majority of deconvolution techniques are $u$-domain methods, probably because of the ease with which Fourier transforms can be calculated either in a computer using the F.F.T. algorithm or optically using the Fourier transforming property of lenses, and also because of the filtering property of the F.T. in the presence of noise. The basis of $u$-domain techniques is equation (7.8), from which

$$H^*(u) = E(u)/G(u) \tag{7.9}$$

where $E(u)$ is the F.T. of $e(x)$. (7.9) shows that, to obtain the unknown spectrum $H(u)$, we must pass the input spectrum, $E(u)$, through a filter with a response of $\frac{1}{G(u)}$.

Stroke (1969), in a discussion of optical deconvolution, notes that in the special case in which $G(u)G^*(u) \approx 1$, $H^*(u)$ is most simply obtained as $E(u)G^*(u)$. In the more general case, the R.H.S. of (7.9) can be obtained optically as $E(u)G^*(u)/|G(u)|^2$. A problem with the F.T. division technique
arises if \( G(u) \) is zero or very small in the range of \( u \) for which \( E(u) \) has appreciable value. Ideally, \( |E(u)| \) should be zero whenever \( |G(u)| \) is zero but in the presence of noise this is unlikely to be the case, and the R.H.S. of (7.9) will be undefined. In certain special cases (Harris, 1966; Swindel, 1970), the \( 1/G(u) \) transfer function of the filter can be modified to take account of the zeros, but in general, when \( G(u) \) has many zeros, deconvolution by F.T. division is not practical. The presence of noise in either or both of \( E(u) \) and \( G(u) \) reduces the accuracy of \( H(u) \) obtained from (7.9). Morrison (1963) uses some useful smoothing techniques on \( E(u) \), before the division is carried out, to reduce the effects of noise. Bracewell (1958), Helstrom (1967) and Schell (1969) obtain expressions for the optimum filter transfer functions to be used in the presence of various errors in \( E(u) \) and \( G(u) \).

All the deconvolution techniques described above require complete knowledge of \( g(x) \) or \( G(u) \). The method described in the next section requires a knowledge of only \( |G(u)| \).

Further, whereas in other techniques small values of \( G(u) \) cause difficulties in the deconvolution procedure, in the new technique even the presence of zeros in \( G(u) \) causes no problems.

7.4.3 Deconvolution Without Phase Information

\( g(x) \), \( h(x) \) and \( e(x) \) are functions of finite extent. Thus, as explained in section 5.1, \( G(w) \), \( H(w) \) and \( E(w) \) are characterized by the positions of their zeros in the complex plane. Denote by \( \{w_{gn}\}, \{w_{hn}\} \) and \( \{w_{en}\} \) the sets of complex zeros of \( G(w) \), \( H(w) \) and \( E(w) \) respectively. It follows from
(7.8) that

\[ \{w_{en}\} = \{w_{gn}\} \cup \{w^*_{hn}\} \]  

(7.10)

If only E(u) is available then it is not possible to identify which members of \{w_{en}\} are included in \{w_{gn}\} and which are included in \{w^*_{hn}\}. But suppose that \|G(u)\| is also known. Then, as explained in section 5.1.3, from \|G(u)\|^2 we can compute \{w_{gn}\} \cup \{w^*_{hn}\}.

By comparing the zeros of E(w) with the zeros obtained from \|G(u)\|^2, the zeros of G(w) can be determined immediately. Thus, compare \{w_{en}\} and \{w_{gn}\} \cup \{w^*_{gn}\} and identify those members (denote them by the set \{w^t_{en}\}) of \{w_{en}\} which are members of \{w_{gn}\} \cup \{w^*_{gn}\}. From (7.10), \{w^t_{en}\} = \{w_{gn}\}. Also, \{w^*_{hn}\} can be obtained by deleting the members of \{w_{gn}\} from \{w_{en}\}. Since \{w_{gn}\} and \{w_{hn}\} are now known G(u) and \|G(u)\|^2 can be determined (except possibly for a linear phase shift) and g(x) and h(x) can be determined unambiguously (except possibly for their absolute positions) given only E(u) and \|G(u)\|. It should be pointed out that if a particular member of \{w_{gn}\} is exactly equal to a member of \{w_{hn}\}, say \(w_{g1} = w_{h1}\), then (7.10) shows that both \(w_{g1}\) and \(w^*_{g1}\) are zeros of E(w).

In this case, it will not be possible to choose all the zeros of G(w) unambiguously, so that more than one g(x), h(x) pair will be possible. However, since in general there is no reason why \{w_{gn}\} and \{w_{hn}\} should have members in common this does not represent a severe limitation of the method.

A detailed deconvolution procedure (Procedure P) using the theory outlined above in this section is now given.
Procedure P
(P.1) Examine $e(x)$ to obtain its extent, $2\beta_e$.

(P.2) Compute $E(u)$ using $e(x) \rightarrow E(u)$. Determine $2M_e$, the number of available samples of $E(u)$ spaced $1/2\beta_e$ apart. For example, if measurements are made only for $|u| < 1/\lambda$, then $M_e$ is the largest integer satisfying $M_e < 2\beta_e/\lambda$. Compute the $2M_e$ members of $\{w_{en}\}$. The $\{w_{en}\}$ are the zeros of the polynomial $P(w)$ defined in (5.10), with $F(u)$, $\beta$ and $M$ replaced by $E(u)$, $\beta_e$ and $M_e$ respectively.

(P.3) Compute $g(x) \cdot g^*(x)$ using $|G(u)|^2 \rightarrow g(x) \cdot g^*(x)$. Note the extent of $g(x) \cdot g^*(x)$ and call it $4\beta_g$. The extent of $g(x)$ is $2\beta_g$. Calculate $2\beta_h$, the extent of $h(x)$, using $\beta_h = \beta_e - \beta_g$.

(P.4) Determine $2M_g$, the number of available samples of $|G(u)|$ spaced $1/2\beta_g$ apart. $2M_g$ is the number of members of $\{\zeta_{gn}\}$, where $\{\zeta_{gn}\}$ is the set of upper-half-plane zeros of $G(w)G^*(w^*)$. Compute the $2M_g$ members of $\{\zeta_{gn}\}$ using the computational method described in section 8.2.1.2.

(P.5) Recall that $\{\zeta_{gn}\} \cup \{r^*_n\} = \{w_{gn}\} \cup \{w_{gn}^*\}$. Compare the members of $\{w_{en}\}$ with the members of $\{\zeta_{gn}\} \cup \{r^*_n\}$ and obtain the $2M_g$ members of $\{w_{gn}\}$ as the subset of $\{\zeta_{gn}\} \cup \{r^*_n\}$ such that $\{w_{gn}\} = \{w_{en}\} \cap ([\{\zeta_{gn}\} \cup \{r^*_n\}]$. The symbols $\cap$ and $\cup$ indicate intersection and union respectively.

(P.6) Obtain $\{w_{gn}^*\}$ as the complement of $\{w_{gn}\}$ with respect to $\{w_{en}\}$.

(P.7) Compute $G(u)$ using (5.8), (5.9) and (5.11) with $F(u)$,
\( \beta, M \) and \( w_n \) replaced by \( G(u) \), \( \beta_g, M_g \) and \( w_{gn} \) respectively.

(P.8) Compute \( \tilde{H}(u) \) using (5.8), (5.9) and (5.11) with \( F(u) \), \( \beta, M \) and \( w_n \) replaced by \( H(u) \), \( \beta_h, M_h \) and \( w_{hn} \) respectively.

(P.9) Compute \( g(x) \) and \( h(x) \) using \( G(u) \rightarrow g(x) \) and \( H(u) \rightarrow h(x) \).

7.4.4 Applications of the Deconvolution Procedure

The theory of the deconvolution procedure given in the previous section is straightforward. However, the associated computations are quite involved, and so it was considered worthwhile trying the procedure on experimental data, contaminated with appreciable amounts of noise. The source, having the distribution \( f(x) \), was composed of two acoustic transducers with aperture distributions \( g(x) \) and \( h(x) \) radiating in air at a wavelength of 1 cm. \( g(x) \) was the 4.2 cm wide composite transducer described as the second example in section (3.4.2). A plan of the transducer is shown in Fig. 3.15 together with its measured aperture distribution. \( h(x) \) was composed of two 1.2 cm diameter circular transducers mounted side by side to give an overall width for \( h(x) \) of 2.8 cm. \( g(x) \) and \( h(x) \) were separated by a distance of 4.6 cm.

Radiation patterns were measured using the procedure of section 3.2.1. Fig. 7.6 shows \( |F(u)| \) and \( |H(u)| \). \( |G(u)| \) is the measured radiation pattern shown in Fig. 3.16.

Fig. (7.7) shows \( r(x) \) computed using \( |F(u)|^2 \rightarrow r(x) \). \( e(x) \) is that part of \( r(x) > V_r \) in the range \( x > 4.6 \) cm. The relatively large value of noise level, \( V_r \), needed to separate \( e(x) \) from the rest of \( r(x) \) indicates that there is considerable noise in \( e(x) \). Since measurements could not be made for
u > 1 cm⁻¹, the values of 2β_e and 2β_g of 7.5 cm and 4.2 cm respectively give values of $M_c$ and $M_g$ of 7 and 4 respectively. Notice that $M_h = 2$ because $2β_e = 2.8$ cm. This means that although there are 14 zeros of $E(w)$ in the visible range $|u| < 1$ there are only 12 zeros of $G(w)$ and $H(w)$ combined, so that two zeros of $E(w)$ will not be zeros of $G(w)$ or $H(w)$ in the visible range. The zeros of $E(w)$ and $G(w)G^*(w)$ are shown in Fig. 7.8 as are the zeros of $H(w)H^*(w)$ which are included as a check. It can be seen that the choice of the members of $\{w_{gn}\}$ is reasonably unambiguous for all eight members. The effect of the relatively large noise level in $e(x)$ can be seen in the gaps between some members of $\{w_{en}\}$ and corresponding members of $\{c_{gn}\} \cup \{c^*_ {gn}\}$. The reconstructed $g(x)$, call it $g_1(x)$, obtained using step (P.9) is shown in Fig. 7.9a together with the actual $g(x)$, call it $g_0(x)$, that was measured as described in section 3.4.3. There is good agreement between $g_0(x)$ and $g_1(x)$. Also shown in Fig. 7.9a are some of the other $g(x)$ closest to $g_0(x)$ that could have been possible if only $|G(u)|^2$ was known. It is clear that the proposed deconvolution technique has allowed us to reconstruct $g(x)$ correctly. Because of the two extra zeros of $E(w)$ that are present, the zeros of $H(w)$ cannot be found simply by removing the members of $\{w_{gn}\}$ from $\{w_{en}\}$. However, if the members of $\{c_{hn}\} \cup \{c^*_ {hn}\}$ are compared with the members of $\{w_{en}\}$, it can be determined whether $c_{hn}$ or $c^*_ {hn}$ is a zero of $H(w)$. Thus, $\{w_{hn}\}$ can be determined. The $h(x)$ reconstructed using these zeros is shown in Fig. 7.9b. This compares very well with the $h(x)$ actually measured using the technique of section 3.4.3, which is also shown in Fig. 7.9b.
The reconstruction procedure described above will allow the brightness temperatures \( g(x) \) and \( h(x) \), of astronomical sources to be found when \( g(x) \) and \( h(x) \) are sufficiently separated and when \( |F(u)|^2 \) and \( |G(u)|^2 \) can be measured. Often, however, it may be impossible to measure \( |G(u)|^2 \) separately from \( |F(u)|^2 \) (for instance, when observing double stars). Even so, there is still a good chance of reconstructing the brightness temperature distributions uniquely (provided their separation is greater than either of their widths). This is demonstrated in the next section.

7.5 Reconstruction of Double, Separated Sources

Consider \( f(x) \), a function of finite extent \( 2\beta \) and its F.T. in the complex \( w \) domain, \( F(w) \). As described in Chapter 5, \( F(w) \) is characterized by the positions of its zeros in the complex plane (call them \( \{w_n\} \)). Define \( f^{(1)}(x) \) to be a function having the same width as \( f(x) \) and let \( F^{(1)}(w) \) be its F.T. in the complex \( w \) plane. Then, if \( |F(u)| = |F^{(1)}(u)| \), it is shown in section 5.1.3 that a suitable expression for \( F^{(1)}(w) \) is

\[
F^{(1)}(w) = F(w)T(w) \tag{7.11}
\]

where

\[
T(w) = \frac{w - w_n^*}{w - w_m^*} \tag{7.12}
\]

In (7.11) and (7.12) \( F^{(1)}(w) \) has been obtained by flipping only one zero of \( F(w) \). Using the convolution theorem for Fourier transforms, (7.11) gives

\[
f^{(1)}(x) = f(x) \ast t(x) \tag{7.13}
\]
where \( t(x) \) is the F.T. of \( T(u) \).

\[
t(x) = \int_{-\infty}^{\infty} \frac{u-w^*_m}{u-w_m} \exp(-j2\pi ux) \, du
\] (7.14)

Assume for convenience that \( w_m \) is in the U.H.P. Then (7.14) can easily be evaluated to give

\[
t(x) = \begin{cases} 
0 & \text{if } x > 0 \\
(x) + j2\pi(w_m-w^*_m)e^{-j2\pi w_m x} & \text{if } x < 0
\end{cases}
\] (7.15)

If \( w_m \) is in the L.H.P., the inequality signs on the R.H.S. of (7.15) are simply reversed. (7.13) and (7.15) show that, in the \( x \) domain, \( f^{(1)}(x) \) is obtained from \( f(x) \) by adding to \( f(x) \) the convolution of \( f(x) \) with a truncated complex exponential. The fact that this convolution does not extend the width of \( f^{(1)}(x) \) outside the range \( |x| > \beta \) can be verified by examining the convolution integral (Hofstetter, 1964). When \( f^{(1)}(x) \) is obtained by flipping a number of zeros of \( F(w) \), \( f(x) \) is successively convolved by the same number of truncated exponentials.

Now consider a case in which \( f(x) \) is composed of two distributions, \( g(x) \) and \( h(x) \) separated by a distance greater than either of their widths. As described in section 5.1.3 and 6.4.3, from a knowledge of only \( |F(u)| \) all the possible \( f^{(n)}(x) \) can be computed. However, one of the effects of the convolution with a complex exponential described in the previous paragraph is to fill in the gap between \( g(x) \) and \( h(x) \) so that it is unlikely that many of the possible distributions, other than the correct \( f(x) \), will consist of two completely separated parts. For those few possible
distributions that are separated into two parts, denote the part to the left by \( g(x) \) and the part to the right by \( h(x) \). From each \( g(x) \) we can compute a \( G(w) \), using \( g(x) \rightarrow G(w) \), which is characterized by its set of complex zeros \( \{w_g\} \). Similarly each \( h(x) \) gives rise to an \( H(w) \) characterized by its set of zeros \( \{w_h\} \). It seems likely, although no mathematical proof has been found, that only the correct \( f(x) \) will give rise to sets \( \{w_g\} \) and \( \{w_h\} \) that are compatible with the set \( \{w_e\} \) (remember that \( \{w_e\} = \{w_g\} \cup \{w_h^*\} \)). This is the basis for the following reconstruction procedure (Procedure Q).

**Procedure Q**

(Q.1) Measure \( |F(u)|^2 \) and compute its F.T. \( r(x) \) using \( |F(u)|^2 \rightarrow r(x) \).

(Q.2) Inspect \( r(x) \). If, after the application of a noise threshold (see section 2.4.4), \( r(x) \) separates into three distinct parts then postulate that \( f(x) \) consists of two separated parts (call them \( g(x) \) and \( h(x) \)). Denote by \( e(x) \) the right hand outer part of \( r(x) \). Denote the widths of \( r(x) \) and \( e(x) \) by \( 4\beta \) and \( 2\beta_c \) respectively. Denote the width of the central part of \( r(x) \) by \( 4\beta_c \).

(Q.3) Compute the sets of zeros \( \{\zeta_n\} \), the zeros of \( F(w)F^*(w^*) \) in the U.H.P., and \( \{w_e\} \), the zeros of \( E(w) \).

(Q.4) Compute all possible \( f^{(m)}(x) \) using the set \( \{\zeta_n\} \), as described in section 6.4.3. Take the few \( f^{(m)}(x) \) (suppose there are \( N \) of them) that are separated into well defined parts and call their separated parts...
g^{(m)}(x) and h^{(m)}(x)$, $1 \leq m \leq N$. Compute the zeros \{w^{(m)}_{gn}\} of $G^{(m)}(w)$, where $g^{(m)}(x) \rightarrow G^{(m)}(w)$. Compute the zeros \{w^{(m)}_{hn}\} of $H^{(m)}(w)$, where $h^{(m)}(x) \rightarrow H^{(m)}(w)$.

(Q.5) Now attempt to choose the correct $f(x)$ from the $N$ possible $f^{(m)}(x)$. The simplest way to do this is to make use of the relationships between $r(x), e(x), h(x)$ and $g(x)$. Denote by $2\beta_{gm}$ and $2\beta_{hm}$ the extents of $g^{(m)}(x)$ and $h^{(m)}(x)$. For the correct $g^{(m)}(x)$ and $h^{(m)}(x)$ the following conditions must hold:

If $\beta_{gm} > \beta_{hm}$, then $\beta_{gm} = \beta_c$ and $\beta_{hm} = \beta_e - \beta_{gm}$ (7.16)

If $\beta_{gm} < \beta_{hm}$, then $\beta_{hm} = \beta_c$ and $\beta_{gm} = \beta_e - \beta_{hm}$ (7.17)

where $\beta_e$ and $\beta_c$ were defined in step Q.2. Test the possible $f^{(m)}(x)$ to see which of them satisfy (7.15) or (7.16). If none of the $f^{(m)}(x)$ satisfy (7.15) or (7.16) then the postulate made in step (Q.2) must be wrong. If only one $f^{(m)}(x)$ satisfies (7.15) or (7.16) this one may be chosen with confidence as the correct one. (Remember, however, that if $f^{(m)}(x)$ is a possible distribution that satisfies (7.16) or (7.17), then $f^{(m)}(-x)$ is also a possible distribution that satisfies (7.16) or (7.17). This ambiguity between a distribution and its mirror image can never be resolved without some phase information.) If more than one $f^{(m)}(x)$ satisfies (7.16) or (7.17) proceed to step (Q.6) to try to resolve the ambiguity.

(Q.6) For the possible $f^{(m)}(x)$ satisfying (7.16) or (7.17) compare \{w^{(m)}_{gn}\} U \{w^{*(m)}_{hn}\} with \{w_{ne}\}. If the sets are
effectively identical for one particular \( f^{(m)}(x) \),

choose this one as the correct distribution.

Although it is not possible to prove that procedure Q will always lead unambiguously to the correct distribution, a computed example illustrates its application for a particular case. Fig. 7.10 shows the postulated brightness temperature, shown as \( f^{(0)}(x) \). Fig. 7.11a shows \( |F(u)| \), the interferogram of \( f^{(0)}(x) \). Since measured interferograms are necessarily truncated at some maximum baseline, \( |F(u)| \) was truncated at the position shown. The zeros \( \{w_n\} \cup \{w^*_n\} \) corresponding to this truncated interferogram intensity are shown in Fig. 7.11b, as are the true zeros of \( F(w) \). \( f^{(1)}(x) \), shown in Fig. 7.10, is the best estimate of \( f^{(0)}(x) \) that can be obtained using the available range of baselines. \( f^{(1)}(x) \) was computed using the members of \( \{w_n\} \cup \{w^*_n\} \) with the correct sign for the imaginary parts. For this example there are 64 possible \( f^{(m)}(x) \). The mirror images of these 64 are also possible.

Only two of these, other than \( f^{(1)}(x) \), are effectively separated into well defined parts. These two are shown in Fig. 7.10 as \( f^{(2)}(x) \) and \( f^{(3)}(x) \). The three \( f^{(n)}(x) \), 

\( 1 < x < 3 \), shown in Fig. 7.10 separate completely into two parts if a noise threshold \( V \) of 5\% is applied. The remaining 61 \( f^{(m)}(x) \) (which we found could be computed and inspected rapidly) all had "noise levels" greater than \( 2V \), and all had appreciable negative parts. It is seen that \( f^{(1)}(x) \) is significantly closer to \( f^{(0)}(x) \) than either \( f^{(2)}(x) \) or \( f^{(3)}(x) \). It is felt that this example gives considerable support for the proposed procedure (Procedure Q) since it was not necessary to choose \( f^{(0)}(x) \) in any special way in order to get the demonstrated result.
FIGURE 7.1. EXAMPLES OF ZEROS OF $G(w)$ AND $R(w)$.

FIGURE 7.2. $F(w)$ MEASURED AND COMPUTED FROM THE ZEROS.
Figure 1.3. Zeros of $G(w)$ and $R(w)$.

Figure 1.4. Radiation pattern of test antenna for example of Section 7.23.
FIGURE 7.5 RECONSTRUCTED DISTRIBUTION, \( f(x) \)

FIGURE 7.6 MEASURED PATTERNS -- \( |F(u)| \), --- \( |H(u)| \)
FIGURE 7.7. $r(x)$, THE FT. OF $|F(w)|^2$. --- MODULUS, ---- PHASE.

FIGURE 7.8. ZEROs OF $G(w)G^*(w^n), H(w)H^*(w), E(w)$

- $\{w_{gn}\} \cup \{w_n\}$
- $\{w_{hn}\} \cup \{w_{hn}\} + \{w_n\}$

MEMBERS OF $\{w_{gn}\}$ MARKED BY ARROW.
FIGURE 7.9a. SOURCE DISTRIBUTION $g(x)$

--- $g_0(x)$  $\times\times\times g_1(x)$  --- $g_2(x)$  $\circ\circ\circ g_3(x)$

FIGURE 7.9b. SOURCE DISTRIBUTION $h(x)$

--- $h_0(x)$  $\times\times\times h_1(x)$
Figure 7.10 Brightness Distributions

\( f_0(x) \) \( \times \times \times f_1(x) \) \( \cdots \) \( f_2(x) \) \( o \cdots o \) \( f_3(x) \).
FIGURE 7.11a. INTERFEROGRAM MODULUS, $|F(u)|$.

FIGURE 7.11b. ZEROS OF $F(w)F^2(w)$ AND $F(w)$.
- $\{W_{fn}\} \cup \{W_{f, n}^2\}$ CALCULATED FROM $|F(u)|^2$, $\triangle \{W_{fn}\}$ CALCULATED FROM $F(u)$. 
CHAPTER 8. Computational Considerations

In this chapter the computational procedures used to obtain the results described in previous chapters are discussed.

8.1 Computation of Fourier Transforms

All Fourier transforms computed in previous chapters were computed using the Fast Fourier transform (F.F.T.) algorithm (Cooley and Tukey, 1965). This method was used because of its very great speed advantage over other methods (Bergland, 1969). Since its introduction in 1965 by Cooley and Tukey, the F.F.T. algorithm has been widely applied to a large variety of engineering and scientific problems (Bergland 1969; IEEE, 1967; IEEE, 1969). Although the F.F.T. algorithm actually computes a Fourier series, it can be used to evaluate approximately a continuous Fourier integral using sampled data. This is an established technique well described in the literature (Bergland, 1969; Cooley et al., 1967). The theory of the method will not be repeated here. Some practical considerations of the application of the technique are briefly discussed below.

Sampling requirements are the most important consideration when the F.F.T. algorithm is used to compute Fourier transforms. The chief requirement is that the data to be transformed should be sampled at a rate equal to or greater than twice the maximum frequency for which its F.T. has significant value. Thus, if $f(x)$ is to be computed using $F(u) \rightarrow f(x)$ and if $f(x) = 0$, $|x| > \beta$, then $F(u)$ must be
sampled at intervals \( \Delta u_f \) such that

\[
\Delta u_f < \frac{1}{2\beta}
\]  \hspace{1cm} (8.1)

Similarly, if \( r(x) \) is to be computed using \( |F(u)|^2 \rightarrow r(x) \), \( |F(u)|^2 \) must be sampled at intervals \( \Delta u_f \) such that

\[
\Delta u_f < \frac{1}{4\beta}
\]  \hspace{1cm} (8.2)

Suppose that \( |F(u)|^2 \) is sampled at intervals of width \( \Delta u \).

Then the maximum value of \( x \) (call it \( x_{\text{max}} \)) for which \( r(x) \) can be computed using the F.F.T. is \( |x_{\text{max}}| = \frac{1}{2\Delta u} \). If \( r(x) \) has significant value for \( |x| > x_{\text{max}} \), that part of \( r(x) \) in the range \( |x| > x \) is folded back into the range \( |x| < x_{\text{max}} \), causing errors in the computed \( r(x) \). This effect is known as "aliasing" (Bergland, 1969). The presence of noise in measured \( F(u) \) or \( |F(u)| \) increases the extent of \( f(x) \) and \( r(x) \) beyond \( 2\beta \) and \( 4\beta \) respectively. For computations on actual measured \( F(u) \) (radiation patterns or interferograms), sampling intervals considerably smaller than those specified in (8.1) and (8.2) are therefore needed to prevent aliasing errors. A sampling spacing smaller than one half that required in (8.1) and (8.2) was used in most of the computations reported in this thesis. In all cases the sampling spacing was sufficiently small so that the value of the transform in the region of \( |x| = x_{\text{max}} \) was negligible (less than 1%), compared with the maximum value of the transform. Whilst the problem of aliasing was being examined, a new transform based on square-wave kernels was developed (Bates, Napier and Chang, 1970). This
transform has advantages over the F.F.T. in situations where aliasing errors are important but where it is not convenient to reduce the sampling spacing. Since this technique was not used to compute any of the results mentioned in the previous chapters, it will not be discussed further here.

Another sampling requirement when using the F.F.T. is that the function to be transformed should be sampled uniformly. Smith (1969) has shown how to estimate the effect of departures from uniform sample spacing, which he calls sampling jitter. Except for the measurement reported in section 4.3.3, all sampling was done manually. In the case of the measured acoustic radiation patterns which were recorded on a pen-recorder as functions of \( \theta \), sampling at uniform spacing in \( u \) required sampling at equal increments of \( \sin \theta \). This was performed by placing a grid calibrated in equal increments of \( \sin \theta \) over the recorded pattern. The patterns were sampled on this grid. Uniform sampling of the astronomical interferograms at uniform increments of baseline was straightforward. The effect of sampling jitter was certainly negligible compared with the effect of measurement noise for the astronomical data used in chapter 6. For the measured radiation patterns, sampling jitter has similar effect to measurement noise and its effect was assumed to be included in the composite noise level described in section 2.4.4.

The final important computational consideration for Fourier transforms is the problem mentioned in section 2.4.2; the effect of transforming a function that is theoretically of infinite extent using a necessarily finite number of
samples. Since functions in the x domain \( f(x) \) or \( r(x) \) are of finite extent, the problem of computing the transform of truncated data only occurs in transforming from the u domain to the x domain. As explained in section 2.4.2, the two major effects of truncation in the u domain are reduction of resolution in the x domain and introduction of spurious ripples in \( f(x) \) or \( r(x) \) near points where they change rapidly. Resolution can only be improved by carrying out data extrapolation which will usually be impracticable. The spurious oscillations in \( f(x) \) or \( r(x) \) can be reduced by weighting \( F(u) \) or \( R(u) \) with data windows that cut-off less abruptly than the rectangular window used in (2.36) or (6.2). The various types of data windows have been described in detail in the literature (Bergland, 1969; Blackman and Tukey, 1958 p. 95). It should be remembered, however, that such data windows further reduce the resolution available in \( f(x) \) or \( r(x) \). In the case of the H.P.M. techniques, satisfactory estimates of \( r(x) \) were obtained without using any data windows. This was because, even for the relatively small antennas (4 wavelengths wide) used in the measurements described in Chapter 3, \(|F(u)|^2\) was small (less than 1% of the peak value) near \(|u| = 1/\lambda\) so that the effect of truncating \(|F(u)|^2\) beyond this point is small. For larger antennas, the effect of truncating the radiation pattern at \(|u| = 1/\lambda\) will be even smaller. No data windows have been used in the computation of the reconstructed brightness temperature distributions for the astronomical cases described in Chapter 6. The accuracy of the reconstructions is determined by the accuracy with which the complex zeros of \( F(w) \) can be computed.
Since accurate computation of the zeros requires $|F(u)|$ to be accurate, a data window should not be applied to $|F(u)|$ during the zero computation and phase reconstruction steps of the reconstruction procedures. If required, smoothing can be applied after the reconstruction has been carried out to reduce the oscillations in $f(x)$ resulting from only a finite number of the zeros being available.

8.2 Computation of Zeros

8.2.1 Computation of Zeros Using Polynomial Method

In section 5.1.2 it is shown that the zeros of $F(w)$ can be obtained as the roots of a polynomial. This technique is now described in detail.

8.2.1.1 Zeros of $F(w)$

The zeros of $F(w)$ were obtained as the roots of the polynomial $P(w)$ defined by (5.10). To calculate the zeros of $P(w)$ it is convenient to express it in the form

$$P(w) = \sum_{n=0}^{2M} a_n w^n$$

(8.4)

It is difficult to obtain an explicit expression for the $a_n$ directly from (5.10). However, (5.10) can be expanded to give

$$P(w) = F(0) \prod_{n=1}^{M} \left( w^2 - \frac{n^2}{4\beta^2} \right) + \sum_{m=1}^{M} \left[ w^2 [F(\frac{m}{2\beta}) + F(-\frac{m}{2\beta})] \right]$$

$$+ w \left( \frac{m}{2\beta} [F(\frac{m}{2\beta}) - F(-\frac{m}{2\beta})] \right) \prod_{n=1}^{M} \left( w^2 - \frac{n^2}{4\beta^2} \right)$$

(8.5)

(8.5) is composed of $M+1$ individual polynomials and can be used to evaluate the $a_n$. The coefficients of the polynomials
\[ \prod_{n=1}^{M} \left( w^2 - \frac{n^2}{4\beta^2} \right) \] can be easily evaluated using a subroutine for multiplying polynomials (IBM 1968, p. 172). Corresponding coefficients of the individual polynomials in (8.5) are then added to give the \( \alpha_n \). It is unlikely that the sampled values of \( F(u) \) that are stored in the computer will correspond exactly to the points \( |u| = m/2\beta \), so an interpolation procedure is needed to give the samples of \( F(u) \) required in (8.5). The order of the interpolation routine used is determined by the spacing between the stored samples of \( F(u) \); for the computations described in Chapters 6 and 7 a second order interpolation method was found to be adequate.

(8.5) is suitable for the most general case in which \( f(x) \) is complex and non-symmetric, and was used to compute the zeros of \( E(w) \) in section 7.4.4. When, as in the case of the astronomical data, \( f(x) \) is real then

\[ F(u) = F^*(-u) \quad (8.6) \]

For the zero computations in Chapter 6, (8.6) can be used to simplify (8.5). In this case even coefficients, \( \alpha_{2n} \), are real and odd coefficients, \( \alpha_{2n+1} \), are imaginary.

The Newton-Raphson method (Korn and Korn, 1968, pg. 719) was used for obtaining the zeros of \( P(w) \) expressed in the form of (8.4). The method can be used to find the real and complex zeros of a polynomial with complex coefficients, as described in Appendix 5.

8.2.1.2 Theory of the Method for Zeros of \( R(w) \)

As explained in section 5.1.3, \( r(x) \) is necessarily of finite extent \( 4\beta \). Then the analysis applied to \( f(x) \) and \( F(w) \) in section 5.1.2 can be applied to \( r(x) \) and \( R(w) \) to give
\[ R(w) = P_R(w) Q_R(w) \quad (8.7) \]

where
\[ Q_R(w) = \frac{\sin(4\pi\beta w)}{2M \prod_{m=-2M}^{w - \frac{m}{4\beta}}} \quad (8.8) \]

and
\[ P_R(w) = \sum_{n=-2M}^{2M} (-1)^n R(\frac{n}{4\beta}) \prod_{m=-2M}^{w - \frac{m}{4\beta}} \quad (8.9) \]

Notice that (8.7), (8.8) and (8.9) are not compatible with the definition \( R(w) = F(w)F^*(w^*) \) when \( F(w) \) is expressed in the form of (5.8), (5.9) and (5.10). The explanation for this is as follows. Suppose that \( f(x) \) can be represented exactly by an exponential Fourier series with a finite number \((2M+1)\) of terms as in (5.5). The autocorrelation \( r(x) = f(x)\overline{f}(x) \) cannot be represented exactly by an exponential Fourier series with \((4M+1)\) terms, but requires an infinite number of terms. The point is that limiting \( f(x) \) to \((2M+1)\) terms of the Fourier series in (5.5) implies that
\[ F(\frac{2m}{4\beta}) = 0, \quad |m| > M, \quad \text{whilst limiting} \quad r(x) \quad \text{to} \quad \text{(4M+1) terms of a Fourier series which leads to (8.7), (8.8) and (8.9), implies that as well as} \quad F(\frac{2m}{4\beta}) = 0, \quad |m| > M, \quad \text{also} \quad F(\frac{2m+1}{4\beta}) = 0, \quad |m| > M. \]

In fact, \( F(\frac{2m+1}{4\beta}) \) is not zero for \( |m| > M \), but since these samples are unavailable because of measurement truncation, (8.7), (8.8) and (8.9) give the best available estimate of \( R(w) \) in the range \(|u| < \frac{M}{2\beta}\). The zeros of \( P_R(w) \), defined in (8.9), are the best available estimates of the zeros of \( R(w) \) in the range \(|u| < M/2\beta \) and are obtained using the method described in section 8.2.1.1 with \( P(w) \) replaced by \( P_R(w) \).
8.2.1.3 Accuracy of the Polynomial Zero Computation Method

The method described in section 8.2.1.1 and Appendix 5 is a simple, rapid and accurate method of computing the zeros of $F(w)$ and $R(w)$. Typical computing time for evaluating the coefficients and solving the polynomial was four seconds for a polynomial of order fourteen. Computation time increases as the square of the order. The zeros of a test polynomial were found to an accuracy of better than five significant figures, which was adequate for operating on measured data whose accuracy is never better than 1%. Several factors prevent the zeros of the polynomial being exactly the zeros of $F(w)$ or $R(w)$; chief of these are: the presence of noise in the samples of $F(u)$ or $R(u)$, incorrect spacing of the samples of $F(u)$ or $R(u)$ and the truncation of $F(u)$ or $R(u)$. All three of these conditions introduce errors into the computed coefficients of the polynomial. The zeros of polynomials can be very sensitive to relatively small changes in the coefficients of the polynomial. Although it is difficult to generalise, as a rule multiple zeros or zeros close together will be most sensitive to small changes in the coefficients (National Physical Lab. 1961, p.59) and complex zeros will usually be more stable than real zeros (Fox and Mayers 1963, p.61).

The problem of computing the zeros of noisy data stems mainly from the fact that, as shown by (8.9) and (5.10), the zeros are determined completely by the values of $F(u)$ or $R(u)$ at isolated sample points. If these points happen to contain relatively large amounts of noise, the computed zeros will be in error. The inaccuracies will be evident when $F(u)$ or $R(u)$
is reconstructed using the zeros. Although the reconstructed $F(u)$ and $R(u)$ will necessarily be equal to the measured data at the sample points, they will not agree at the points in between. A method of improving the accuracy of the zeros by making use of the measured data at points between the sampling points is described in section 8.2.1.4.

The spacing between the samples of $F(u)$, $R(u)$ or $E(u)$ is determined by the estimated widths of $f(x)$, $r(x)$ and $e(x)$, respectively. Slightly different estimates of these widths will result in slightly different samples of $F(u)$, $R(u)$ or $E(u)$ which in turn alter the coefficients of the polynomials. It is important that the computed zeros are not too sensitive to these changes since there is always uncertainty in estimated widths. It has been found that the most important zeros, the ones on the real axis and those nearest the origin of the complex plane, are relatively insensitive to uncertainties in the sampling width. The sensitivity increases with increasing distance from the origin of the complex plane. As an example, consider the computation of the zeros of $E(w)$ in the measurement described in section 7.4.4. Estimates of $2\beta_e$, the width of $e(x)$, that are compatible with Fig. 7.7 are $2\beta_e = 7.5$ cm and $2\beta_e = 7.15$ cm. This represents a variation in width of approximately 5%. Corresponding polynomial coefficients computed using samples of $E(u)$ for these two estimates of $2\beta_e$ differ by as much as 50% for particular coefficients. However, the zeros computed for the two values of $2\beta_e$ are shown in Fig. 8.1, and it is clear that only those zeros well off the real axis have been affected. Notice, in particular, that no zeros have crossed over the real axis; this is
important when phase information is to be inferred from the sign of the imaginary parts of the zeros, as in the techniques of sections 6.3.2, 7.2.2 and 7.4.3.

The effect of truncation of $F(u)$ or $R(u)$ on computed zeros decreases as the real part of the zero decreases. This is expected since the missing samples of $F(u)$ or $R(u)$ have reduced effect as $|u|$ decreases. The effect of truncation was found typically to prevent multiple zeros appearing exactly as multiple zeros and to perturb the positions of zeros nearest $u = u_{\text{max}}$. Tests on ideal data that had been generated within the computer showed that only the zero nearest $u_{\text{max}}$ was significantly affected by truncation, but it did not cross the real axis. This zero could usually be corrected using the zero refinement procedure described in section 8.2.1.4.

8.2.1.4 Zero Refinement Procedure

If $f(x)$ were exactly zero for $|x| > \beta$ and if it could be represented exactly by a Fourier series with a finite number of terms, then the zeros computed from the polynomial defined by the samples $F(\frac{m}{2B})$ would be correct and the $F(u)$ reconstructed from these zeros would agree with the measured $F(u)$ at all values of $u$. In practice, however, noise and measurement truncation ensure that $F(u)$ is not exactly band-limited and the $F(u)$ reconstructed from the polynomial zeros will not be in exact agreement with the measured $F(u)$. It has been found that the positions of zeros can be improved by using a non-linear optimization technique that refines the positions of the zeros computed from $P(w)$ or $P_w(w)$, to provide the best least-squares fit between the actual measured $F(u)$
or \( R(u) \) and the \( F(u) \) or \( R(u) \) reconstructed from the zeros. Non linear optimization procedures have previously been found valuable for computing zeros of polynomials (Hanson, 1971). When the zeros of \( F(u) \) were being computed, the zero positions were refined by minimizing the function

\[
\varepsilon = \int_{-u_{\max}}^{u_{\max}} |B(u)|^2 du
\]  

where

\[
B(u) = R_m(u) - P(u)Q(u)
\]  

When the zeros of \( R(u) \) were being refined, \( B(u) \) in (8.10) was replaced by \( A(u) \), where

\[
A(u) = |F_m(u)| - |P(u)Q(u)|
\]  

\( F_m(u) \) is the measured data and \( P(u) \) and \( Q(u) \) are defined in (5.9) and (5.11). The integral in (8.10) was approximated by a summation of \( B(u) \) over all sampled values of \( u \). Thus the function actually minimized was \( \sum_{n=-N}^{N} |B(n\Delta u)|^2 \), where \( \Delta u \) is the spacing between samples of \( F(u) \) or \( |F(u)| \) stored in the computer and \( N = \frac{u_{\max}}{\Delta u} \). \( \Delta u \) was usually approximately a quarter of the Fourier sample spacing.

The method of Rosenbrock (1960) was used to carry out the minimization. The computer program implementing Rosenbrock's method, modified to increase its efficiency, was written by Chan (1970). Rosenbrock's method is a first order method for successively varying a set of parameters in order to minimize a function of the parameters. Although it is a first order method and does not require gradient calculations, it exhibits an approximate second order convergence.
The bulk of the execution time is spent on evaluating the function to be minimized, so it is essential that the function evaluation routine be programmed as efficiently as possible. The function to be minimized is $\varepsilon$ defined in (8.10) and (8.11) or (8.12). A point to be remembered in computing $\varepsilon$ is that, for the astronomical data for which (8.6) holds and the zeros are symmetrical about the imaginary axis, the integral in (8.10) need only be evaluated within the limits $0$ and $u_{\text{max}}$. Also, since $Q(u)$ is independent of the positions of the zeros of $P(w)$, $Q(u)$ need not be continually re-evaluated as the positions of the zeros are varied.

The zero refinement routine using Rozenbrock's method was found to be very effective. The interferogram reconstructed from the zeros of the polynomial often differed from the actual measured interferogram by as much as 20% of its peak value for isolated values of $u$. The interferogram reconstructed from the zeros after they had been refined seldom differed from the measured interferogram by more than 2% of the peak value over the whole range of $|u| < u_{\text{max}}$. It was noticed that the refinement procedure had a smaller effect on the zeros of $F_m(u)$ than on the zeros of $R_m(u)$. This is because the zeros calculated from the polynomial for $F_m(u)$ are constrained by both $|F_m(u)|$ and $\arg F_m(u)$, whilst the zeros of $R_m(u)$ are constrained by only $|F_m(u)|$. The computing time taken by the refinement procedure to optimize the positions of five zeros was forty-five seconds, the time increasing approximately in proportion to the square of the number of zeros. An example will demonstrate the effectiveness of the zero refinement procedure. Consider the zeros of $R(w)$ for
Example 2 in section 6.3.2. Fig. 8.3 shows six zeros of $R(w)$ computed from the polynomial $P_R(w)$ defined in (8.9). Errors are evident in these zeros because any real zeros of $R(w)$ should be double zeros. Assume that the correct zeros are the zeros obtained from the polynomial in the positions $(637 + j320)$, $(1181 + j0)$ and $(1518 + j0)$. $|F(u)|$ reconstructed from these zeros is shown in Fig. 8.2 and it can be seen that it does not agree very well with the measured $|F(u)|$. However, when the positions of these zeros are refined to give the best least-squares fit between the measured and reconstructed $|F(u)|$, the more accurate zeros in Fig. 8.3 are obtained. $|F(u)|$ reconstructed from these refined zeros is shown in Fig. 8.2 and agrees well with the measured $|F(u)|$.

8.2.2 Other Zero Computation Procedures

The polynomial method, combined with the zero refinement, is simple and direct and should be used whenever possible to compute the zeros of $F(w)$ or $R(w)$. For the measurement of section 7.2.3, however, in which the constant reference signal can be considered as equivalent to introducing a delta function into $f(x)$, the number of terms required in a Fourier series representation of $f(x)$ becomes very large and the polynomial method of computing the zeros of $F(w)$ is not suitable. A different method for computing the zeros of $F(w)$ or $R(w)$ is therefore needed for this case.

Since $R(w)$ is the analytic continuation of $R(u)$ into the complex $w$ plane, it satisfies

$$R(w) = \int_{-\infty}^{\infty} r(x) \exp(-j2\pi wx)dx$$

(8.13)

where (8.13) is the analytic continuation of the inverse of
(2.12). (8.13) can be used to compute $R(w)$ over the whole complex plane. Notice that

$$R(w) = R(u+jv) = \int_{-2\beta}^{2\beta} [r(x) \exp(2\pi vx)] \exp(-j 2\pi ux) \, dx \quad (8.14)$$

$r(x)$ can be easily obtained using $|F(u)|^2 \rightarrow r(x)$. (8.14) shows that by weighting $r(x)$ with the function $\exp(2\pi vx)$ and Fourier transforming the product, the value of $R(w)$ for any value of the imaginary variable $jv$ can be found. Computationally, only relatively small values of $|v|$ may be handled easily because of the rapid growth of $\exp(2\pi vx)$. For the measurement described in section 7.2.3, zeros of $R(w)$ were expected in the region $v=2$. Since for this measurement $r(x)$ existed in the range $|x| < 4$, for $v=2$ the weighting function ranges between $10^{-20}$ and $10^{20}$. This range is too great for (8.14) to be computed with sufficient accuracy to locate zeros of $R(w)$. (8.14) also shows that because of the exponential weighting the edges of $r(x)$ are emphasised. Near its edges $r(x)$ is usually small and is likely to be contaminated by noise, further reducing the likelihood of finding zeros well off the real axis.

The zeros of $R(w)$ were located by searching the complex plane using (8.14) to evaluate $R(w)$. If (8.14) is evaluated for a fixed value of $v$ using the F.F.T., the value of $R(w)$ is found at points along a line parallel to the real axis. The value of $R(w)$ was computed in this way along lines of constant $v$ in the U.H.P. The lines were separated by increments $\Delta v = 0.001$, so that the positions of the zeros could be located to within $\pm 0.0005$. The zeros of $R(w)$ for $|u| < 8$
were located by searching the complex plane in this way. The
zeros were found to be regularly spaced for |u| > 6 so that
the positions of the zeros for |u| > 8 were estimated by
direct extrapolation. The zeros in this range did not need
to be computed very accurately since \( F(u) \) was only recon-
structed in the range |u| < 1. The reconstruction of \( F(u) \)
was carried out using all zeros for which |u| < 20.

The method of searching the complete complex plane in
this way was relatively slow (approximately 20 minutes com-
puting to search the required area of the upper half of the
complex plane) because \( R(w) \) was computed at values of u and v
well away from its zeros as well as at values close to its
zeros.

A more efficient alternative to searching the whole of
the complex plane is to make use of the fact, mentioned in
section 5.1.1, that an entire function falls smoothly to its
zeros and has no minimum points. This allows the zeros to be
found iteratively and \( F(w) \) or \( R(w) \) need only be computed at
points in the complex plane lying near the path taken from
the first estimate of the zero to its actual position. A
simple but effective routine based on this method was devised.
The value of \( R(w) \) was computed at \( w = w_n \), an initial starting
estimate of a zero, and at the four points \( w_n \pm \Delta u, w_n \pm j \Delta v \).
The increments \( \Delta u \) and \( \Delta v \) were chosen to be approximately \( \frac{1}{16\beta} \),
a quarter of the Fourier sample spacing for \( R(u) \). Whichever
of these four points gave the smallest value of \( |R(w)| \) was
chosen as the next estimate of \( w_n \). If none of the four points
gave \( |R(w)| \) smaller than \( |R(w_n)| \), \( \Delta u \) and \( \Delta v \) were halved. The
process was repeated until $\Delta u$ and $\Delta v$ were less than some prescribed amount (usually 0.001 of the absolute value of the zero position) at which stage the current value of $w_n$ was taken as the zero position. To use this simple search routine it is necessary to have a suitable expression for $R(w)$. As pointed out at the start of this section, $r(x)$ cannot be represented as a Fourier series, because of the presence of the $\delta(x)$ term. However, if the delta function is separated out, the rest of $r(x)$ can be represented as a Fourier series. Thus, if the constant reference signal is of strength $S$, $r(x)$ can be written as

$$r(x) = S\delta(x) + \sum_{m=-2M}^{2M} \frac{[R\left(\frac{m}{4\beta}\right) - S]}{4\beta} \exp(j\pi mx/2\beta)$$  (8.15)

which, when substituted into (3.14) gives

$$R(w) = S + \frac{\sin\left(\frac{4\pi\beta w}{4\pi}\right)}{4\pi\beta} \sum_{m=-2M}^{2M} \frac{(-1)^m(R\left(\frac{m}{4\beta}\right) - S)}{(w - m/4\beta)}$$  (8.16)

(8.16) is a suitable expression for evaluating $R(w)$ in the search routine. This method gave the same zero positions as the F.T.T. method described at the start of this section, and did not locate any zeros in the region $v \approx 2.0$ for the reason explained above. It was considerably faster than the F.T. method (approximately one second computing time per zero). The simple search routine was also used for finding the zeros of ideal (computer generated) functions for which an analytic expression for $F(w)$ was available.
8.3 Reconstruction of $F(u)$ and $f(x)$ from the Zeros of $F(w)$

Once the zeros of $F(w)$ had been found, $F(u)$ was reconstructed in both modulus and phase using (5.8), (5.9) and (5.11). This procedure is computationally straightforward although it is worthwhile noting that in the computation of $F(u)$ at the sample points $u = \frac{n}{2\beta}$, which are needed for the reconstruction of $f(x)$, (5.9) simplifies to

$$Q(\frac{n}{2\beta}) = \frac{(-1)^{M-n}(2\beta)^{2M-1}}{\pi(M-n)! (M+n)!}$$ \hspace{1cm} (8.17)

$f(x)$ was reconstructed using (5.5) and (5.7). The most efficient way of evaluating (5.5) is of course to use the F.F.T. algorithm. However, when $f(x)$ is real (as for astronomical data) it is just as simple, when $M$ is small, to make use of the fact that $F(\frac{n}{2\beta}) = F^*(-\frac{n}{2\beta})$, in which case (5.5) simplifies to

$$f(x) = F(0) + 2 \sum_{n=1}^{M} \text{Re}[F(\frac{n}{2\beta})] \cos(\pi nx/\beta) + \text{Im}[F(\frac{n}{2\beta})] \sin(\pi nx/\beta)$$ \hspace{1cm} (8.18)

In section 6.4.3 a procedure was introduced for obtaining all the possible $f(x)$ that could have given rise to a particular measured $|F(u)|$. The procedure involved "flipping" the zeros of $F(w)$. If $L$ zeros are to be flipped then there are $2^L$ possible $f(x)$ that are defined by all the possible combinations of these $L$ zeros. It is essential that $L$ be kept as small as possible if the number of $f(x)$ to be printed out and examined is to be kept small enough to be practical. To this end it is worthwhile making some comments on which zeros are important in the flipping procedure, and which zeros can be expected to have negligible effect on $f(x)$.
when they are flipped.

Recall, (5.8), that \( F(u) \) can be represented as the product of the two factors \( P(u) \) and \( Q(u) \). When a zero is flipped, only \( P(u) \) is altered. Consider only cases for which \( f(x) \) is real, so that (6.12) holds. Suppose that a given combination of zeros gives \( F(0)(u) \). Then if the \( m^{th} \) zero \( w_m = u_m + jv_m \) is flipped, \( F(0)(u) \) is changed to \( F(1)(u) \) where (6.12) shows that

\[
F(1)(u) = F(0)(u) \frac{u^2 - u_m^2 - v_m^2 - j2uv_m}{u^2 - u_m^2 - v_m^2 + jv_m} \quad (8.19)
\]

It is only necessary to consider the effect of flipping the zero on the value of \( F(u) \) at its sample points. The theory of section 5.1.2 shows that each sample point except \( F(0) \) has associated with it a zero. In general \( u_m = m/2\beta \).

Examination of (8.19) shows that if the imaginary part of \( w_m \) is much smaller than a Fourier samplespacing, that is if

\[
v_m \ll 1/2\beta, \quad (8.20)
\]

then flipping \( w_m \) only affects \( F(m/2\beta) \), and leaves other samples essentially unchanged. Flipping \( w_m \) has its greatest effect on \( F(m/2\beta) \) when it is closest to the sample point, that is when \( u_m = m/2\beta \). In this case, (8.19) gives

\[
F(1)(m/2\beta) = F(0)(m/2\beta) \frac{v_m^2 + jmv_m/\beta}{v_m^2 - jmv_m/\beta} \quad (8.21)
\]

(8.21) shows that if (8.20) holds and if the real part of \( w_m \) is much greater than the imaginary part, then the only effect of flipping \( w_m \) is to invert the \( m^{th} \) term in (8.18). The importance of this, with respect to the average value of \( f(x) \),
can be determined by comparing $|F(m/2\beta)|$ with $|F(0)|$. If $|F(m/2\beta)| \ll |F(0)|$ it can be expected that the effect on $f(x)$ of flipping $w_m$ will be negligible.

If $w_m$ is so far removed from the real axis that

$$v_m \gg M/2\beta$$  \hspace{1cm} (8.22)

where $M$ is the highest order present on the R.H.S. of (8.18), then substituting (8.22) into (8.19) shows that flipping $w_m$ has negligible effect on $f(x)$.

Unfortunately, most zeros satisfy neither (8.20) nor (8.22). The only very general statement that can be made is that the relative importance of flipping a zero can be judged approximately by the relative size of its associated sample of $F(u)$, with respect to $F(0)$. In this respect, it should be noted that if a certain combination of zeros gives rise to a wholly positive $f(x)$, flipping those zeros that are associated with the largest samples of $F(u)$ is likely to give an $f(x)$ that contains significant negativeness.
FIGURE 8.1. ZEROS OF $E(w)$ FOR TWO DIFFERENT VALUES.
OF $\beta e \circ 2\beta e = 7.15 \text{cm}$, $\triangle 2\beta e = 7.5 \text{cm}$. 
Figure 8.2: Modulus of Interferogram. --- Measured Mod.
Reconstructed, --- from polynomial zeros, o o o from refined zeros.

Figure 8.3: Computed Zeros. △ polynomial zeros
○ refined zeros
CHAPTER 9. Conclusion

9.1 Suggestions for Future Research on H.P.M. Techniques

In Chapters 2, 3 and 4 several new techniques have been described that allow the modulus and phase of an antenna's aperture distribution and/or the modulus and phase of its radiation pattern to be inferred from the modulus of a measured radiation pattern. It has been demonstrated, using acoustic experiments, that these techniques are of practical value. Although the techniques are useful for measuring acoustic antennas, it is expected that they will be most useful for making measurements on electromagnetic antennas. Due to unavailability of suitable radio frequency equipment none of the techniques has as yet been used to measure an electromagnetic antenna. Demonstration of the usefulness of the methods in electromagnetic measurements is obviously necessary if the techniques are to become used by antenna engineers. In this respect it is worth noting that some of the techniques will be more easily applied electromagnetically than acoustically. For example, continuously variable phase shifters are more readily available at microwave frequencies than they are at acoustic frequencies.

Little has been said about the actual interpretation of the equivalent one-dimensional aperture distributions that are obtained from the H.P.M. techniques. The equivalent one-dimensional distribution for a linear array can supply considerable information, as was demonstrated in section 3.3.2. The interpretation of the equivalent distribution of, for example, a large irregularly shaped aperture in which the field is not linearly polarized would not be straightforward.
This point would need investigation for each particular electromagnetic application.

Recall equation (2.13) which, with equation (2.4), can be considered to be the basic equation for the H.P.M. techniques. The essential difference between each of the various H.P.M. techniques is the way in which each technique allows the required cross correlation term, $h(x)\hat{\otimes}g^*(x)$ or $g(x)\hat{\otimes}h^*(x)$, to be separated from the unwanted autocorrelation terms $g(x)\hat{\otimes}g^*(x)$ and $h(x)\hat{\otimes}h^*(x)$. In section 3.3.4, a pattern multiplication method of Blum et al. (1961) was mentioned that has the advantage that the unwanted autocorrelation terms never appear. This corresponds to a suggestion made by Deschamps (1967). As noted in section 3.3.4, the multiplication technique is useful when radioastronomical sources are used to measure radiation patterns. It is possible that the technique could also be useful for conventional pattern measurements whenever the provision of an accurate multiplier is more convenient than the accurately phase controlled reference signals required by some of the H.P.M. techniques. Experimental investigation of this point would be worthwhile.

9.2 Suggestions for Future Research on Complex Zero Techniques

The results obtained in Chapter 6 demonstrate that the procedures described are practical and useful methods for identifying and removing phase errors in measured astronomical interferograms. To date, the techniques have been used only for processing data in one dimension. Using the concept of aperture synthesis (Jennison, 1966), interferometers are currently being widely used to obtain two dimensional (2-D) maps of the celestial sphere. For the techniques of Chapter 6
to become widely applicable, therefore, it is necessary that they be used to process 2-D data.

The two procedures (K and L) for determining whether or not significant phase errors are present in measured data are directly applicable in 2-D situations. Denote by \([x_1,x_2]\) and \([u_1,u_2]\) pairs of orthogonal coordinates on, respectively, the celestial sphere and the interferometer baseline plane. Denote by \(F_m(u_1,u_2)\) the measured 2-D interferogram. The 2-D brightness temperature distribution \(f_m(x_1,x_2)\) and its autocorrelation function \(r_m(x_1,x_2)\) are obtained as the 2-D Fourier transforms of \(F_m(u_1,u_2)\) and \(|F_m(u_1,u_2)|^2\) respectively. Then, following the principle of Procedure K, the extents of \(f_m(x_1,x_2)\) and \(r_m(x_1,x_2)\) can be compared to see if phase errors in \(F_m(u_1,u_2)\) are indicated. Following the principle of Procedure L, \(F_m(u_1,u_2)\) and \(|F_m(u_1,u_2)|^2\) can be weighted by 2-D triangular gate functions and their 2-D Fourier transforms examined for negativeness. The phase correction procedure (Procedure M) could be applied to 2-D data by correcting the phase of \(F_m(u_1,u_2)\) along closely spaced projections on the \(u_1-u_2\) plane of great circles on the celestial sphere. The number of projections needed to give a significant reduction in the errors in \(f_m(x_1,x_2)\) would require careful investigation as the amount of computation time required to repeat Procedure M many times could become prohibitive.

In the work reported in Chapters 5, 6 and 7, the properties of entire functions of order one (exponential type) have been applied to formulate phase reconstruction procedures Bates\(^*\) has suggested that the properties of entire functions

\(^*\)R.H.T. Bates - private communication
of order two could also be of value in the solution of "phase problems". Entire functions of order two are entire in both the u and x domains so that the advantages to be gained by characterizing a function by its zeros could be exploited in both domains. This suggestion could provide a worthwhile area for further research.

The new deconvolution procedure using complex zeros that is presented in Chapter 7 has not been tested on a wide variety of problems. This needs to be done to prove the usefulness of the technique. It is worth emphasizing that the procedure is more widely applicable than deconvolution by division of Fourier transforms which has become the conventional method. The new procedure can be used whenever the division procedure is applicable.
APPENDIX 1 - Equipment Used for Acoustic Experiments

Oscillator: Advance S.G.66 Low Frequency signal generator, chosen because it can provide the relatively high voltage (100 volts peak to peak) needed to drive the electrostatic transducers. Frequency stability was better than 0.02% over the time taken to measure a radiation pattern.

Receiving antenna and Amplifier: Bruel and Kjoer ¼ inch diameter condenser microphone coupled to a Bruel and Kjoer Audio Frequency Spectrometer Type 2112. The spectrometer was operated in the narrow bandwidth mode with the input filter centred at 31.5 KHz and a bandwidth of 7.3 KHz.

Linear detector: This was an integrated-circuit detector constructed by Martin (1969). A calibration test showed it to be linear to 0.1% over a dynamic range of 40 db, which was adequate for the acoustic experiments.

Pen recorder: Hewlett Packard X-Y Recorder, type 7035B. Linearity 0.1%.

Phase Meter: This was constructed by the author using an established design principle (Johnson, 1969). It was calibrated using the calibrated phase shifter on a Hewlett Packard 3300A Function Generator with 3304A sweep/offset plug-in. The output was proportional to the phase difference between the two input channels to an accuracy of better than 3 degrees over a dynamic range of the input signal of approximately 30 db.

Phase Shifter: The linear phase shift needed for the experiment reported in section 3.4.4 was obtained by moving a probe microphone along a length of matched acoustic slotted line. The movement of the probe was controlled by a servomotor fed by the output from the diode function generator. The acoustic slotted line was simply constructed from a 30 cm length of
0.65 cm diameter circular brass tubing with a 0.15 cm wide, 20 cm long slot milled in the wall. Matching was achieved by packing the end of the waveguide with acoustically absorbing material.
APPENDIX 2 - Comparison of Extents of $r_m(x)$ and $f_m(x)$

If measurement distortion has negligible effect on $F_m(u)$ then equations (6.2) and (6.3) give

$$F_m(u) = F(u) \exp(j\psi(u)) \quad \text{(A.1)}$$

Define $\psi(x)$ by $\exp(j\psi(u)) \rightarrow \psi(x)$ so that

$$f_m(x) = f(x) * \psi(x) \quad \text{(A.2)}$$

So, unless $\psi(u)$ is small, or has the special form $\arg T_n(u)$ for a particular value of $n \ll 2^L$ (see section 5.1.3 for definitions of $T_n(u)$ and $L$), then the convolution in (A.2) shows that the extent $2\beta_m$ of $f_m(x)$ will be significantly greater than $2\beta$. Since $\psi(u)$ will usually be random there is only a very small probability that $\psi(u)$ is not small and yet has one of the special forms ($\arg T_n(u)$) that make $\beta_m \approx \beta$. Note that theorems given by Sheehan (1962) show that $\beta_m < \beta$ is impossible if $\psi(u)$ is the only significant distortion of the measured interferogram. This demonstrates the validity of the assertion, made in step (K.3), that the inequality (6.5) holds if $\gamma > 1$. The assertions that $N(u)$ and $\psi(u)$ are small if $|\gamma| \ll 1$ and that $N(u)$ and/or truncation distortion are more significant than $\psi(u)$ if $\gamma < -1$, are corollaries of the above reasoning.
Define \( \square(x) \) and \( \Lambda(x) \) by \( \text{rect}(u/2u_{\text{max}}) \rightarrow \square(x) \) and \( \text{tri}(u/2u_{\text{max}}) \rightarrow \Lambda(x) \) respectively and note that

\[
\square(x) = 2 \frac{u}{u_{\text{max}}} \sin(2\pi xu_{\text{max}}) \tag{A.3}
\]
\[
\Lambda(x) = \frac{u}{u_{\text{max}}} \sin^{2}(\pi xu_{\text{max}}) \tag{A.4}
\]

where \( \sin(x) \) is defined in the glossary. Suppose that \( N(u) \) and \( \Psi(u) \) are negligible, so that equations (6.2) and (6.3) give

\[
F_{m}(u) = F(u) \text{rect}(u/2u_{\text{max}}); \quad R_{m}(u) = R(u) \text{rect}(u/2u_{\text{max}}) \tag{A.5}
\]

and equation (6.9), and the subsequent definition of \( r_{m}(x) \), give

\[
\hat{F}_{m}(u) = F(u) \text{tri}(u/2u_{\text{max}}); \quad \hat{R}_{m}(u) = R(u) \text{tri}(u/2u_{\text{max}}) \tag{A.6}
\]

from which it follows that

\[
f_{m}(x) = f(x) \star \square(x); \quad r_{m}(x) = r(x) \star \Lambda(x) \tag{A.7}
\]
\[
\hat{f}_{m}(x) = f(x) \star \Lambda(x); \quad \hat{r}_{m}(x) = r(x) \star \Lambda(x) \tag{A.8}
\]

Now \( \Lambda(x) \) is non-negative. Since both \( f(x) \) and \( r(x) \) are non-negative, if \( N(u) \) is negligible then \( f_{m}(x) \) is non-negative whereas if \( \Psi(u) \) is appreciable then \( f_{m}(x) \) need not be non-negative. Conversely, \( r_{m}(x) \) need not be non-negative, even if \( N(u) \) is negligible, because \( \square(x) \) oscillates between positive and negative values. Finally, if in a particular case it is found that either or both of \( f_{m}(x) \) and \( r_{m}(x) \) have
negative parts, and both \( \hat{f}_m(x) \) and \( \hat{r}_m(x) \) are non-negative, then it must be concluded that both \( N(u) \) and \( \psi(u) \) are negligible compared with truncation distortion.

It is worth mentioning that \( \text{tri}(u/2u_{\max}) \) is not the only gate function that could have been used in Procedure L. Any gate function with a wholly positive F.T. could have been used, which implies that any gate function that is an autocorrelation function would be suitable. For example, the autocorrelation functions defined by Schell (1963) could also be used. \( \text{tri}(u/2u_{\max}) \) was chosen for its simplicity.
The crux of Procedure M is step (M.11), the theoretical basis of which is now given. Neglecting \( N(u) \) and truncation distortion

\[
F_m(u) = F(u) \exp(j\Psi(u))
\]  \hspace{1cm} (A.9)

In practice, \( F_m(u) \) is observed only for a limited number of discrete values of \( u \) denoted by the finite set \( \{h_n\} \), so that \( \Psi(u) \) can be thought of as a function having the value \( \Psi(h_n) \) at each point \( h_n \) and being absolutely continuous everywhere on the \( x \) axis. Consequently, \( \Psi(u) \) possesses an analytic continuation \( \Psi(w) \) into the complex plane, so that

\[
F_m(w) = F(w) \exp(j\Psi(w))
\]  \hspace{1cm} (A.10)

Denote by \( P \) and \( Q \) positive real numbers for which, in computations such as those required in steps (M.3) and (M.9), \( P \) is so large that \( PF(w)_n \) will not compute to zero, and \( Q \) is so small that \( QF(w) \) will always compute to zero for any point in \( D \), which denotes the region of the complex plane which is searched for zeros in steps (M.3) and (M.9). In general, there will be finite sets of points \( \{\sigma_n \in D\} \) and \( \{\nu_n \in D\} \) for which \( |\exp(j\Psi(w))| > P \) in a neighbourhood of \( \sigma_n \) and \( |\exp(j\Psi(w))| < Q \) in a neighbourhood of \( \nu_n \). If no member of \( \{\sigma_n\} \) is close to any member of \( \{w_n\} \) then all the zeros of \( F(w) \) are zeros of \( F_m(w) \). However, in general \( F_m(w) \) has more complex zeros than \( F(w) \), because

\[
F(\nu_n) = F(w_k) = 0
\]  \hspace{1cm} (A.11)
which is the theoretical justification for $K$, obtained in step (M.8) and used in step (M.9) being greater than $M$, obtained in step (M.2) and used in step (M.3). Under these conditions, and provided the measurement distortion is not great enough to produce the errors described in the following paragraph, step (M.11) is always successful. In fact, the application of Procedure C to the examples shows that step (M.11) can be successful when the phase distortion is appreciable.

The zeros which are operated on in step (M.11) are not $\{w_k\} \cup \{v_n\}$ but $w_{mrk}$. Because of measurement distortion the subset $\{w'_{mrk}\}$ of $\{w_{mrk}\}$ will only be an approximation to $\{w_k\}$. Since $\{w_{mrk}\}$ is computed from data containing $\Psi(u)$, it must be assumed that $\{v_n\}$ is a proper subset of $\{w_{mrk}\}$ so that, if measurement distortion is sufficiently severe, one or more members of $\{v_n\}$ may displace approximate zeros of $\Psi(w)$ in $\{w'_{mrk}\}$, in which case (refer to section 5.1.3), instead of getting an estimate of $\{w_k\}$, we get an estimate of $\{\tau_k w_k + (1-\tau_k)w^*_k\}$, for some particular member (call it the $n$th member) of the $2^L$ possible sequences $\{\tau_k\}$ (see section 5.1.3). Then step (M.12) instead of giving an estimate of $f(x)$, gives an estimate of $f^{(m)}(x)$ (defined in section (5.1.3)) which is a good approximation to a brightness temperature compatible with $R(u)$ but not with $\Psi(u)$. If one or more members of $\{v_n\}$ are close to members of $\{w_k\}$ then $\{w_{mrk}\}$ is short of some members it should have for step (M.11) to be successful, and some of the members of $\{w_{mrk}\}$ are again on the wrong side of the real axis. If phase distortion is so severe that the members of $\{w_{mrk}\}$ are very closely spaced then $f_\tau(x)$ obtained
in step (M.12) will be a good approximation to one of the $2^L$ possible $f^{(m)}(x)$, with all possible values of $m$ being equally likely. This is just as it should be because it corresponds to such severe phase distortion that $\arg F_m(u)$ is virtually random, in which case there can be no hope of choosing the correct $\{w_{rk}\}$. 
APPENDIX 5 - Newton-Raphson Technique for a Complex Polynomial

The analysis given here is a generalization to complex polynomials of a Newton-Raphson method (IBM 1968, Pg. 181) for finding the roots of a real polynomial. If the complex coefficients, \( a_n \), of the polynomial have real and imaginary parts \( a_n \) and \( b_n \) respectively, (8.4) can be written

\[
P(w) = \sum_{n=0}^{M} (a_n + jb_n)w^n
\]

(A.12)

Let \( w = u+jv \) be a starting value for a root of \( P(w) \). Then

\[
w^n = (u+jv)^n
\]

(A.13)

Define \( u_n \) as the real terms of expanded equation (A.13) and \( v_n \) as the imaginary terms. Then for \( n = 0 \), \( u_0 = 1.0 \), \( v_0 = 0.0 \) and for \( n > 0 \)

\[
u_n = u \cdot u_{n-1} - v \cdot v_{n-1}
\]

(A.14)

\[
v_n = u \cdot v_{n-1} + v \cdot u_{n-1}
\]

Let \( X \) be the real terms of (A.1) and \( Y \) the imaginary terms. Then

\[
X = a_0 + \sum_{n=0}^{M} a_n u_n - b_n v_n
\]

(A.15)

\[
Y = b_0 + \sum_{n=0}^{M} a_n v_n + b_n u_n
\]

and, from the definitions of \( u_n \) and \( v_n \),
\[
\frac{\partial X}{\partial u} = \sum_{n=1}^{M} n a_n u^{n-1} - n b_n v^{n-1} \\
\frac{\partial X}{\partial v} = -\sum_{n=1}^{M} n a_n v^{n-1} + n b_n u^{n-1} 
\] (A.16)

Notice that equations (A.14), (A.15) and (A.16) can be evaluated iteratively for \( n = 1 \) to \( M \) by storing \( u_{n-1} \) and \( v_{n-1} \). Now, the distance, \( \Delta w = \Delta u + j \Delta v \), to the next estimate of the zero of \( P(w) \) is given by the Newton-Raphson method (Korn and Korn 1968, p.719) as

\[
\Delta w = -\frac{P(w)}{P'(w)} 
\] (A.17)

where \( P'(w) \) is the differential of \( P(w) \), so that, after applying the Cauchy-Riemann equations (Morse and Feshback 1953, p.357)

\[
\Delta u = \left( Y \frac{\partial X}{\partial v} - X \frac{\partial X}{\partial u} \right) \sqrt{\left( \frac{\partial X}{\partial u} \right)^2 + \left( \frac{\partial X}{\partial v} \right)^2}
\] (A.18)

\[
\Delta v = -\left( X \frac{\partial X}{\partial v} + Y \frac{\partial X}{\partial u} \right) \sqrt{\left( \frac{\partial X}{\partial u} \right)^2 + \left( \frac{\partial X}{\partial v} \right)^2}
\]

(A.18) is evaluated using (A.15) and (A.16). For the next iteration \( u = u + \Delta u \), \( v = v + \Delta v \). The procedure was programmed by making the necessary changes (including the \( b_n \) in (A.15) and (A.16)) to an existing program, POLRT (IBM 1968, P.181) that solved real polynomials. As each zero was found, the polynomial was reduced by dividing out the zero, but final iterations for all zeros were carried out on the original polynomial to avoid accumulated errors in the reduced polynomial. Computations were carried out in double-precision arithmetic.
REFERENCES


JENNISON, R. C. (1958), "A Phase Sensitive Interferometer Technique for the Measurement of the Fourier Transforms of Spatial Brightness Distributions of Small Angular Extent", ...


LEVIN, B.Ja. (1964), Zeros of Entire Functions, American Mathematical Society, Providence, R.I.


National Physical Laboratory (1961), Modern Computing Methods, Her Majesty's Stationery Office, London.


