N O T I C E

1 JULY 1981

MULTIFREQUENCY APERTURE-SYNTHESIZING MICROWAVE RADIOMETER SYSTEM (MFASMR)

HUGHES PROPRIETARY CONTRACT NAS 5-25755

VOLUME I

NASA CR- 166792

HUGHES

HUGHES AIRCRAFT COMPANY

PREFACE

Part I, Sections 1 through 10, and Part II, Section 1, contain a large body of background and explanatory material. Part of this material is necessary to an understanding of the analysis of the MFASMR system, which is presented in Part II, Section 2, in response to Task 1 of the work statement.

The result of the analysis in Part II, Section 3, indicates that the MFASMR system does not exhibit high performance because much of the available thermal power is not used in the construction of the image and because the image which can be formed has a resolution of only ten lines. Because of this unfavorable result, Task 2, which called for a comparison with filled aperture systems in terms of weight, complexity, reliability, deployment, storage, performance, cost and development risk, became unnecessary. Therefore, the study effort was redirected to emphasize the analysis of image reconstruction and an understanding of aperture synthesis.

Task 3, image reconstruction, or, more formally, the image quality computer study, was performed in three ways: 1) two-dimensional inverse Fast Fourier Transformation, 2) Wiener filtering plus the inverse Fast Fourier Transformation and 3) maximum entropy reconstruction. The results of these computer analyses are summarized in Part II, Section 2, and are presented in detail in Volume II, Appendix. In addition, the computer programs are delivered as part of the contract effort.

Task 4, comparison with conventional aperture synthesis systems, has been completed. This effort is reported in Part I, Section 8, and in Volume II, Appendices. The appendices in Volume II also contain discussions in more detail of the theory of radiometry by correlation and aperture synthesis.

Because referencing is implicit in the nature of this material, a bibliography for Part I is in Section 11, Part I, and a bibliography for Part II is in Section 6, Part II. In the appendices in Volume II, the references are footnoted within each appendix.

CONTENTS

VOLUME I
PART I - ANALYTIC STUDY AND HARDWARE DESCRIPTION

Page

PART II - COMPUTER STUDY OF IMAGE RECONSTRUCTION WITH MFASMR DATA

VOLUME II - APPENDIX

ILLUSTRATIONS

PART I

ILLUSTRATIONS – PART II

TABLES

PART I

PART II

PART I

Analytic Study and Hardware Description

1. INTRODUCTION

1.1 RECAP OF THE STUDY OF THE MFASMR SYSTEM

This report describes a study of a device for high resolution radiometric mapping. This Multifrequency Aperture-Synthesizing Microwave Radiometer System (MFASMR) is an example of a class of imaging systems which form images by a two-step process. They offer an alternative imaging method which competes with more familiar imagers such as lenses and pencil-beam scanning antennas.

Direct imagers sense the individual resolvable elements in the imaged scene (object space) and build up the image by recording the intensity of each resolvable element (pixel). Lenses display all the pixel intensities simultaneously on a surface called the image plane. In photography all intensities are recorded simultaneously on light-sensitive film. In the microwave region the most successful direct imaging method has been the use of scanning, pencil-beam antenna. The antenna beam looks at the scene pixels, one-at-a-time. The antenna output power is measured by a microwave receiver. Then a display device, such as a CRT, forms image pixels with brightness proportional to the receiver output. They are ordered to match the corresponding pixels in the original scene.

Such image-making processes are familiar due to long usage. The MFASMR and related devices form images in a manner which is not so familiar. In addition to the image plane, there is a second surface in optical systems called the frequency plane or Fourier plane. A set of measurements in this plane can be used as the input to a Fourier transform operation. The transform output is the image. MFASMR and similar systems make such measurements in the Fourier plane in order to form images.

In a little while it will be shown that the electromagnetic flux in the frequency plane is the two-dimensional spectrum of the image-plane field. This spectrum is a spatial frequency spectrum. Each spectrums are not as well known as more familiar temporal frequency spectrums. However both spectrum types possess many analogous properties.

There are several methods of making the frequency plane measurements. The most common method involves movement of an antenna over the Fourier plane while cross-correlating its output with the output of a stationary reference antenna. This process is used by many radio telescopes. It is called aperture synthesis. Alternately the antenna pair can remain stationary and the operating frequency varied in order to obtain the required samples. MFASMR uses this approach.

In the case of aperture synthesis, the collection of samples is usually divided into two steps. The first step consists of movement of the roving antenna along a track. The fixed and roving antennas are at the ends of a variable baseline. The result can be thought of as an interferometer with variable antenna spacing. After the roving antenna reaches a maximum distance from the fixed antenna, the baseline is rotated to a new orientation. The roving antenna repeats its excursion along the new baseline. Eventually, the baseline rotates 180°. If then coincides with its original orientation, but with baseline ends interchanged. Since the antennas are identical, the interchange of baseline ends does not affect measurement. Therefore, the measurements during the first 180° rotation can be used for the measurement values over the next 180° rotation. Thus, samples over a circular disk in the Fourier plane are obtained. In common parlance, only the measurement process which occurs as the roving antenna moves along the baseline is called aperture synthesis. The second step, consisting of rotation of the baseline, is called rotation synthesis or supersynthesis.

Now MFASMR does not use rotation synthesis. Instead, it uses antennas at the ends of two fixed baselines at right angles to each other. Also, MFASMR cannot operate at a set of frequencies dense enough to meet the Nyquist criterion. Only frequencies where there are no transmitters can be used. Thermal radiation is very weak compared to man-made sources. Such sources easily destroy the radiometer accuracy. Only bands reserved for radiometers, or bands where transmitters are not operating for other reasons, can be used. The MFASMR design identifies five such spectrum windows at 2.4, 4.75, 9.5, 19.0 and 38.0 GHz. Both the undersampling and poor distribution of samples damage the data collected.

Some of the damage done to the data can be repaired, if we are willing to use more general methods of forming the image from the spectrum samples than merely using the Fourier transform. These processes of forming images from measurement data are called image reconstruction.

Compensation for undersampling the spectrum depends on the fact that the image signal is redundant. Because of this redundancy, missing data can be replaced with values predicted from knowledge of sample values around the missing value. The average difference between this predicted value and the true value can be made much less than the average difference to be expected when redundancy is lacking.

If the restoration process is linear, there are analytic methods called rate distortion theory for computing the rms image distortion as a function of the amount of data lost in the process of sampling and coding the image data. This rate distortion theory has been used in this study to compute the percentage of image distortion to be expected in MFASMR systems with plausible parameters. It is found that about 30% image distortion is to be expected. This distortion is too great to render this approach attractive.

However, this result does not settle the question as to whether MFASMR is a viable system or not. A significant restriction was placed on the image reconstruction processes when rate distortion theory was used to determine system performance. Namely it was postulated that image reconstruction was a linear process. In fact, reconstruction need not be linear.

During the last decade, a group of nonlinear data processing techniques have been developing which are called by names such as "maximum entropy method" or "maximum likelihood method."

These methods are recursive. Multiple cycles around the recursion loop maximize the output data entropy or other closely allied parameters. Signal statistics do not need to be known a priori, as in linear reconstruction methods. Data statistics are taken into account by their influence on the recursion process.

An excellent review of these methods is provided by an IEEE Press Reprint Volume, Modern Spectrum Analysis, edited by Donald Childers. Spectrum estimation, image reconstruction and adaptive arrays are treated as aspects of the maximum entropy method (MEM) and its relatives.

MEM processing has been controversial, in spite of the excellent results often obtained. Although it often outperforms linear methods, the MEM recursion cycle has turned out to be unstable in many cases. The causes of instability have not been well understood. In addition, the maximization of entropy, or its twin, information, has appeared to be a very abstract objective compared to such down-to-earth objectives as minimization of rms data distortion.

However, Stephen Warnecke (References 1, 2, 3, 4) has developed a MEM algorithm for image reconstruction from frequency plane samples which has constraints built into the recursion loop which ensure stability. It has yielded excellent results compared to linear reconstruction methods. Dr. Wernecke has kindly given us a FORTRAN program using his MEM algorithm. The results obtained using this MEM algorithm to reconstruct images from simulated MFASMR data are given in Volume II of this study report.

The simulated MFASMR data is produced by use of microwave radiometer image signals provided by Robert Moore, Head of the Microwave Radiometer Group at NOTS, Inyokern. These images were generated by a Ka-Band pencil-beam, scanning radiometer. A Fourier-transform operation turns these images into spatial frequency spectrums. These spectrums can now be sampled in the same manner as MFASMR samples the actual spectrum. These sample sets, as well as sample sets corresponding to other frequency-plane imagers, provide the data output for the computer study of MEM image reconstruction reported in Volume II.

1.2   STUDY EXTENSION TO THE PROPERTIES OF FOURIER-PLANE IMAGERS

There are several themes which will be developed in this study which show why frequency-plane imagers are superior to direct imagers for orbital mapping. At this point, these major themes will be sketched in broad outline, permitting the reader to see the goals of the detailed system analysis as it develops.

The first and most important theme is a direct extension of the MFASMR approach, namely, the conviction that frequency-plane imagers will be the dominant type of radiometric orbital mapper in the future. This conviction stems from this study of frequency-plane imagers.

Later in this report, the nature of the frequency plane will be sharpened up. However, the most important property of the frequency plane needs no further analysis. The frequency plane is the electromagnetic field, generated by the scene being viewed, at the surface of imaging lens or reflector, which acts as a decoder of this frequency-plane data. Decoding transforms the data into a form we can understand. This form is an image of the scene.

Building better hardware for the purpose of decoding the frequency-plane data impinging on optical systems has preoccupied people since the time of Newton.

In the last few decades a new idea has grown up. This new strategy for forming images involves the two-step process mentioned in Section 1.1. Instead of transforming the frequency plane data into an image by means of optical hardware, the frequency plane data is measured by means of small probes moving over this plane. Once this data has been recorded, the function of the imaging lens or reflector is carried out by a computer program.

Neither the probes which measure the frequency-plane nor the computer are anywhere near as large as the hardware they supplant. This is especially true when the hardware supplanted is a very large scanning antenna constructed to do high resolution radiometric mapping from orbit. The replacement of the imaging antenna by a computer algorithm has many other advantages besides small size. The computer program is highly flexible. The hardware is not.

We can construct programs for transformation from frequency-plane data to image with many different properties. We can transform in such a way as to make the image the most faithful replica of the scene. Or we can transform so as to obtain the best visual appearance.

If we please, we can measure the properties of the image obtained and then change the transform computation so as to make the image more and more like some standard defining a "best" image. This is comparable to adaptive arrays. However, we do not need to make physical changes in a hardware imager, such as the setting of phase shifters. Our adaptive process is only a recursive loop in a computer program.

Indeed, we can transform the same data in many different ways in parallel computations in order to output images of the same scene optimized in many different ways simultaneously.

This approach focuses attention on the transform processes which have replaced the imaging antenna. As a result, our theme of replacing hardware imagers by software imagers causes heavy emphasis to be placed on these transform processes. These processes are coming to be known as "image reconstruction."

The second theme is the observation that although frequency-plane imagers use small antennas compared to direct imagers of comparable resolution, their sensitivity and ability to resolve temperature variations in the mapped region is just as good as the direct imagers.

Roughly speaking, the antennas of Fourier imagers are sized to see the entire field of view subtended by the region being mapped. This selection of antenna beamwidth is dictated by the consideration that these antennas measure periodicities in the brightness of the scene which extend throughout the entire mapped region. The magnitude and phase of all possible periodicities comprise the spatial frequency spectrum of the scene. It is this spectrum which is Fourier-transformed to form the final image. The scanning antenna of the competing direct imager will be sized to resolve a single pixel in the scene. If the final image has M by M resolvable lines, the scanner will have an area M^2 greater than the area of the Fourier imager antennas.

Now, it is an interesting fact that the large scanner antenna and the little frequency-plane antennas have the same total capture cross section. That is, their capture cross sections integrated over spheres surrounding the antennas are the same. A fundamental antenna theorem states that if the antenna is matched to its transmission line feeder, this total cross section is always λ^2 when the operating wavelength is λ (Reference 5). It is easy to see that, as a consequence of the above theorem, the average power received by the scanner antenna and the frequency sampling antenna is the same.

In order to prove this assertion, let the thermal flux density from the scene be ρ (watts/m²-steradian); let the solid angle subtended by the field-of-view be Ω. If there are M^2 pixels in the scene, the solid angle subtended by the scanning beam of the direct imager, $\Delta\Omega$, is equal to Ω/M^2. The frequency plane antenna is sized to reduce its cross section outside of its field-of-view (FOV) to zero. Therefore, its average cross section must satisfy the relation

$$\overline{A}_{FPA}\Omega = \lambda^2 \quad \text{or} \quad \overline{A}_{FPA} = \frac{\lambda^2}{\Omega} \tag{I-1-1}$$

The thermal power received by the frequency plane antenna is

$$P_{FPA} = \text{(flux density)(FOV solid angle)(average cross section)}$$

$$= \rho\Omega\left(\frac{\lambda^2}{\Omega}\right) = \rho\lambda^2 \text{ watts}$$

In the case of the scanner, ideally the antenna cross section outside the beam is reduced to zero. Therefore

$$\overline{A}_{BEAM}\Delta\Omega = \lambda^2 \quad \text{or} \quad \overline{A}_{BEAM} = \frac{\lambda^2}{\Delta\Omega} \tag{I-1-2}$$

so the power received by the beam is

$$P_{BEAM} = \rho \, \Delta\Omega \left(\frac{\lambda^2}{\Delta\Omega}\right) = \rho\lambda^2 \tag{I-1-3}$$

Therefore, the power received by the two antennas is the same, namely $\rho\lambda^2$.

If we concentrate our attention on a single pixel, we see that it delivers a power of $\rho\lambda^2/M^2$ to the frequency-plane antenna. The same pixel delivers power to the scanner in pulses. When the scanner beam falls on the pixel, it delivers power to the scanner equal to P_{BEAM} or $\rho\lambda^2$. However, the antenna must scan all of the other M^2 pixels, so it spends only $(1/M^2)$ of the

total time looking at a single pixel. So the average power delivered by the pixel to the antenna is the duty factor times the power level during the pulse. This is $(1/M^2)(\rho\lambda^2) = \rho\lambda^2/M^2$. So the average power delivered by a pixel to the sensor is the same in both cases.

Since aperture synthesis uses two antennas, the fixed reference and the roving antenna, which see the entire FOV, the frequency-plane imager actually receives twice the power received by the equivalent direct imager. However, the number of frequency plane samples required for input to the Fourier transform operation is twice the number of pixels produced. Usually, the sample doubling occurs because M^2 frequency plane samples are taken which are complex numbers. The noise power of the real and imaginary parts of the sample add to double the noise in the M^2 pixel-bin outputs from the transform. Since both noise power and signal power are doubled, the signal-to-noise ratio (SNR) of the frequency-plane imager is the same as the SNR of the direct imager.

Another source of confusion lurks in the radio-telescope literature. Hewish (Reference 6) asserts that an aperture-synthesis device has a sensitivity which is reduced by the factor, (number of positions of the movable element)$^{-1/2}$, as compared with a conventional antenna of the same resolving power. For proof he refers to an earlier paper by himself and Ryle (Reference 7).

Without attempting to find the earlier obscure reference, it is easy to see that the factor arises because Hewish is considering a point source. In that case, the Fourier plane samples add incoherently (power-wise) while the scanner antenna adds the point source signal inputs to the array coherently (amplitude-wise). This is true if the scene is only one point source, such as a radio star. However, if the scene is a surface which radiates noncoherently, the thermal radiation from all the surface elements will add power-wise in both the aperture-synthesis device and the direct imaging antenna. This is true even when the scene is a set of point sources which are not coherent with respect to each other. For this case, Hewish's factor disappears. This observation is proven by the detailed derivations of the performance of various imager types given as appendices to this report. Thus Hewish's observation is true, but applies only to the trivial case of a scene consisting of a single point source.

Scanning the literature will show that the correct treatment of coherency conditions for frequency-plane imagers caused much trouble before these conditions were rigorously defined. Hence, it is not surprising to find observations like the one by Hewish in the early papers on Fourier imagers. This theme will be worked out by detailed sensitivity calculations.

A third theme, which will be of importance in this paper, is the observation that rotation synthesis is a technique which will be used independently from aperture synthesis. Usually rotation synthesis is thought of as only part of the total aperture-synthesis process, rotation of the interferometer baseline being needed in order to make measurements over an area. However, rotation synthesis is a technique which stands alone and is very useful even when aperture synthesis is not used.

As Kraus demonstrates (Reference 8), the multiplying interferometer measures the visibility function, $V_0(S_\lambda) e^{j2\pi S_\lambda \Delta\phi}$, as the baseline length in wavelengths, S_λ, is varied from zero to some maximum value D. This function will be identified as the spatial frequency function of the scene at a later time.

Let this function be measured over a set of baseline orientations, θ, extending over 180° in angular displacement. Let the set of orientations be dense enough to make the measurements of V_0 meet the Nyquist condition. Then it can be shown that the two-dimensional Fourier transform of V_0 is an image of the scene in the region selected by the interferometer antennas.

In order to have a compact notation, we will indicate a Fourier-transform pair of functions like this:

$$G(f) \rightleftharpoons P(x) \qquad \text{OR THIS} \qquad G(f) = F\{P(x)\} \qquad (I-1-4)$$

where $G(f)$ is a spectrum of some kind and $P(x)$ is a function of a linear or angular displacement. Then the assertion in the last paragraph that the image is the Fourier transform of the two-dimensional visibility function can be written in the form

$$V(S_\lambda, \theta) \overset{2-d}{\rightleftharpoons} B(\phi, \theta) \qquad (I-1-5)$$

where ϕ is the nadir angle and $B(\phi, \theta)$ is the radiometric brightness of scene.

A valuable feature of this process is the fact that all the baseline orientations can be measured in time succession yet no sensitivity is lost, if the interferometer antennas are correctly sized to receive power from a field-of-view defining the scene which is to be mapped. This is in strong contrast to systems where $B(\phi, \theta)$ is mapped by a pencil-beam scanner.

The scanning array or reflector covers an aperture area as shown in Figure 1-1. The comparable frequency-plane probes are shown to the right of the scanner aperture.

The total array or reflector must be present all the time. If the aperture is "thinned," flux from the scene streams through the gaps created by thinning. If the aperture is thinned enough to produce a significant weight saving, sensitivity is also significantly reduced. The thinned array is of course just as big as the original antenna. Therefore, thinn-ng is of no help in stowage and deployment.
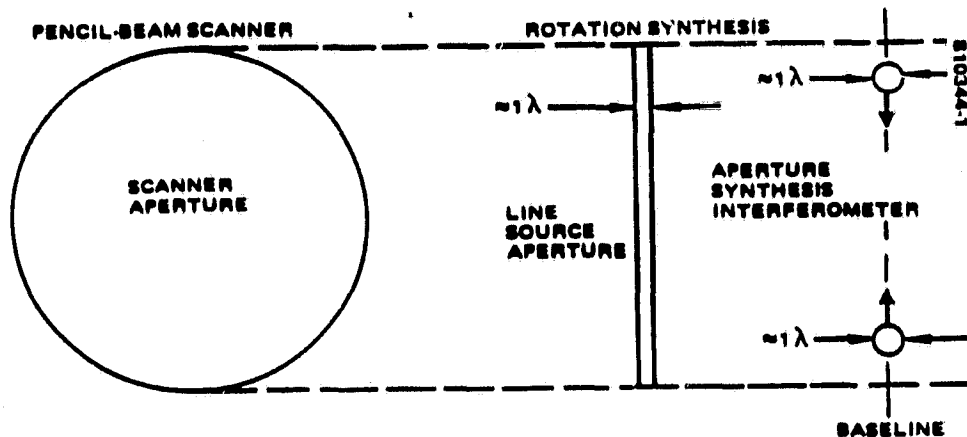
FIGURE 1-1. PENCIL-BEAM SCANNER CONTRASTED WITH ROTATION
SYNTHESIS IMAGER HAVING THE SAME SENSITIVITY

Up to now we have described the measurement of $V(S_\lambda, \theta)$ by aperture synthesis.

However, Kraus (Reference 8) in his Eq. (6-83) gives the relation

$$V(S_\lambda) \rightleftharpoons B(\psi) \qquad\qquad (I-1-6)$$

where $B(\psi)$ is the projection of the source brightness on the axis defined by the baseline oriented at angle θ. That is $B(\psi)$ is the scene brightness as a function of the semiapex angle of the conical antenna patterns of the linear array. These patterns define hyperbolas on the surface defining the scene. For small FOV, Eq. (I-1-6) is obvious.

Now $B(\psi)$ can be measured by any line-source antenna. Then $V(S_\lambda)$ can be computed in a data processor and stored. This can be done for a set of line source orientations identical to the baseline orientations used in the aperture-synthesis case. This is the same $V(S_\lambda, \theta)$ generated by aperture synthesis. The two-dimensional transform indicated in Eq. (I-1-5) again produces the image.

Thus we have the option of using a line source, instead of aperture synthesis, to measure $V(S_\lambda, \theta)$ and then form an image from this function.

This can be a valuable option. Aperture synthesis involves the movement of one interferometer antenna over an area in the Fourier plane. This involves quite a bit of mechanical complexity.

In addition, aperture-synthesis is limited in bandwidth by the requirement that the outputs of the two interferometer antennas must correlate. Some line sources do not have this restriction. This can be of great importance when high resolution and good sensitivity are desired at the same time.

A final theme is emphasis on coherency. Frequency plane measurements are coherency measurements. Therefore, a correlating radiometer is needed to make these measurements. A correlating radiometer is not a radiometer mode but one of two possible radiometer types. The other radiometer type measures power, rather than coherency. Either radiometer can operate in one of many possible modes, such as Dicke, total-power, phase-switching, etc. So the division of radiometers into coherent and noncoherent (power-measuring) types is comparable to the classification of radars as either coherent or noncoherent.

As in the radar case, the theory of coherent radiometers is more subtle. However, the capabilities of correlating radiometers are much greater than the capabilities of radiometers which only measure power. It should not be thought that because the theory is more complicated, the equipment is also more complicated. It is often mentioned in the literature that the correlator is more complicated because it requires two receiver channels. However, in most remote sensing applications data about both signal polarizations are required. Two noncoherent radiometers (two receiver channels) are needed to measure both polarizations. Now both polarizations can pass through a single correlating radiometer by injecting them into the correlator in phase quadrature. So again, two polarization measurements require two channels.

2. BASIC CORRELATING RADIOMETER CONFIGURATIONS

2.1 COHERENCE MEASUREMENT

Before looking at frequency-plane imagers in detail, it is necessary to review the characteristics of correlating radiometers. They differ from ordinary radiometers in that they measure coherence or correlation between two channels. The ordinary radiometer measures power in a single channel. In many ways, these radiometer types mirror the classification of radars into noncoherent (power-measuring) and coherent radars.

Correlating radiometers operate in many of the same modes as noncoherent radiometers. For instance in both cases the simplest mode of operation is as a total power radiometer. The correlating version is shown in Figure 2-1.
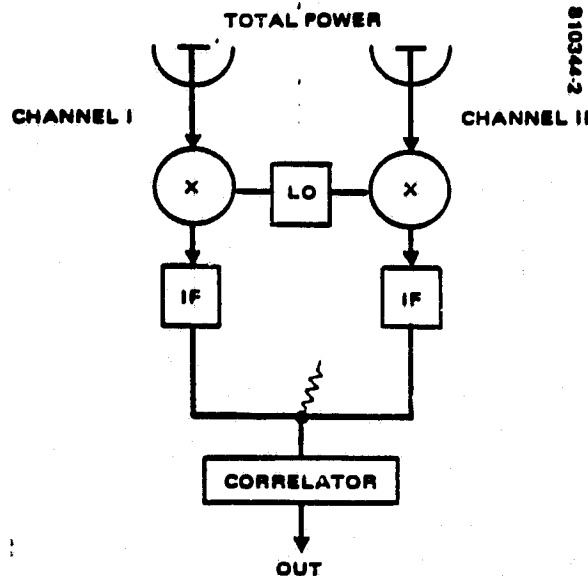
FIGURE 2-1. TOTAL POWER CORRELATING RADIOMETER

The input to this radiometer configuration is shown as two antennas. This is the way the antenna is used is aperture synthesis. One antenna is fixed at the end of the interferometer baseline. The other antenna is the roving antenna which moves along the baseline. Of course, when the maximum baseline length is long, the two channels are not fed directly from the antennas with transmission line. Instead, the signals are reduced to IF, amplified, and distributed.

In the case of Fourier-transform spectrometers, very often there is only one antenna. This single signal is then divided in a hybrid. The hybrid outputs go to the mixers. However, a preferred configuration is still the two-antenna configuration. In this dual configuration there is no portion of the signal channels which is common to both channels. Therefore, none of the radiated thermal power associated with the channel losses appears in the output as part of the signal. Since the radiation comes from different lossy elements, it does not correlate and is rejected.

This makes radiometer calibration much simpler. The effect of internal radiation and of reflected radiation due to channel mismatches is very complicated. A paper by Otoshi (Reference 9) bears out this point.

A similar problem, which is even harder to control, is due to mixer noise. Due to the passage of bias current through the mixer diodes, they act as noise sources. The effective source temperature is ibout 400°K. This radiation travels out through the microwave channels. A portion is reflected at mismatches such as hybrids, circulator, bends, and the antenna itself. Although the radiation is noise, it is coherent with itself. Therefore, a standing wave is formed. If a mismatch produces total reflection, the peak of the mixer-noise standing wave will be $(1 + 1)^2$ x 400°K = 1600°K. The measured mixer-noise error is proportional to a voltage at the radiometer output. It can be seen that if 0.16°K is allotted for mixer-noise error, the reflected mixer noise must be reduced by 10^4 (-80 db). This assumes that the standing wave loop can be anywhere in the microwave channel. If its position can be stabilized in the channel, the errors produced can be calculated out. However, this can be hard to do, especially in wideband systems.

Figure 2-2 shows the effect of mixer noise on a National Radio Astronomy Observatory radio telescope. The long-period fluctuation is due to mismatch at the antenna input. The short-period fluctuation with about 10°K amplitude is due to reflection off the edge of the Cassegrain subreflector.

The dual antenna setup makes it easier to obtain the required 80 db isolation. Mixer noise reflected back to the mixer making the noise finds nothing to be correlated with in the other channel and is eliminated by whatever means the radiometer uses to eliminate other kinds of receiver noise. Only mixer noise which couples from one channel to the other will be treated like a signal and appear in the output. Careful design of the dual antennas can ensure that the coupling between them is very small. When isolators are also used in the two channels, the required 80 db isolation is obtainable.
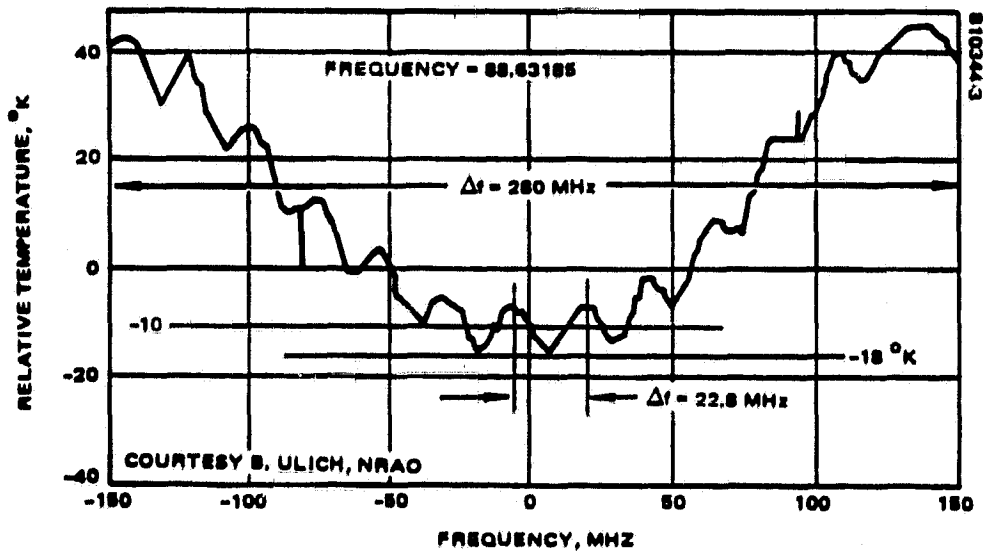
.FIGURE 2-2. EFFECTS OF MIXER NOISE ON MEASURED ANTENNA
TEMPERATURE

In the case of a pencil-beam scanning antenna, no increase in thermal
power results from using two antennas. Both antennas must produce beams
which see the same pixel at the same time. For instance, in the case of a
reflector, the two antennas would be produced by using a dual feedhorn like
that used for a monopulse feed. However, in the case of aperture synthesis,
each antenna may see the total FOV without robbing power from the other
antenna. Therefore, a two-to-one gain in signal power is obtained for this
case.

Total power radiometers do not modulate the input signal. As a result,
fluctuation of the receiver gain produces fluctuation in receiver noise which
cannot be distinguished from the signal. For instance, let the system noise
temperature be 2000°K, let the gain stability be one percent ($\Delta G/G = 10^{-2}$),
then the output receiver noise fluctuations will be 20°K rms. This is an
order of magnitude greater than typical temperature resolutions.

Engineers here at Hughes Space and Communications Group have achieved
remarkable gain stabilities. As a consequence, they have been using the total
power mode. However, the conventional attitude has been that it is too hard
to get the required stability under field conditions. Therefore, the usual
design approach has been to modulate the thermal noise before it becomes
mixed with receiver noise. The modulation then permits separation of the
signal from the receiver noise in the video portion of the radiometer.

The original radiometer of this type was the Dicke. It had an RF switch which switched the receiver input between the antenna and a load. The temperature difference between load and antenna then appears as a thermal signal which is square wave modulated. After square-law detection, the square wave becomes a voltage wave. Synchronous demodulation by the RF switch actuation waveform then picks the thermal difference signal out of the receiver noise.

The corresponding correlation radiometer is the phase-switched radiometer shown in Figure 2-3 on the following page.

Both of the radiometer modes just discussed are not fully coherent. The output of the correlator is at zero frequency. Let the input to Channel 1 be $[A(t) + jB(t)]$. The input to Channel 2 is $[A(t + \tau) + jB(t + \tau)]$. The correlator multiplies these waveforms together and averages the products to give the correlation function in the form $[\overline{A(t)\ A(t + \tau)} - \overline{B(t)\ B(t + \tau)}]_1 + 2j[\overline{A(t)\ B(t + \tau)} + \overline{A(t + \tau)\ B(t)}]_2$. It is the modulus of this final expression which emerges from the radiometers we have been discussing. We cannot get at the real component of the correlation, $[\ ]_1$, nor the imaginary component, $[\ ]_2$. In the same manner, the in-phase and quadrature receiver noise powers add power-wise in the output.
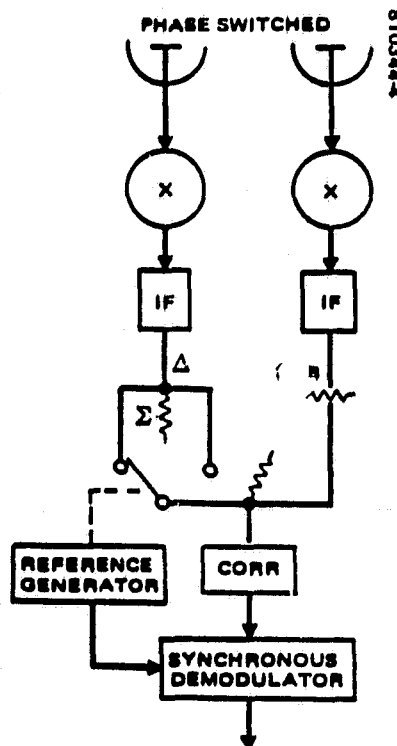
FIGURE 2-3. PHASE-SWITCHED
CORRELATING RADIOMETER

This same series of operations on the signal occur in a homodyne radar, except that one channel is replaced by a sinusoidal reference. Both systems are degenerate in the quantum mechanical sense. Two separate signal modes, corresponding to two eigenvalues have assumed the same values and have become indistinguishable.

In order to avoid this defect, the two radiometer channels can be offset in frequency by using two local oscillators. These LOs are frequency offset by means of a servo loop or by feeding one LO signal through a phase-shifter. This shifter keeps increasing the phase change at a steady rate in order to produce the frequency offset.

The output of the correlator consists of DC terms such as $\overline{A^2(t)}$, which are removed and the cross product terms given above. This signal is demodulated by two phase detectors. One phase detector is a switch whose switching waveform or reference is in-phase with the frequency offset. The other phase detector reference is in quadrature with the offset. The in-phase detector selects only [ ]$_1$. The quadrature detector sees only [ ]$_2$.

A block diagram for this frequency-offset mode is shown in Figure 2-4 below.

Most modern radiotelescopes use this frequency offset technique (References 10, 11). Older instruments use the phase-switching method. They recover the in-phase and quadrature components by shifting the signal from one end of the baseline 90° with respect to the signal from the other end of the baseline. Then they perform a second homodyne detection. These two correlator outputs are often called "promp" and "semi-promp."

Wesseling (Reference 12) is one of the first to point out the substantial advantages of the fully-coherent radiometer over the earlier circuits.

Figure 2-5 shows the microwave portion of a correlating radiometer. The signal enters through the hybrid at the center of the assembly. The two channels are at the top right and bottom left. At the extreme right a noise diode mount can be seen. The diode output is injected into both channels in a servoed bridge arrangement for measuring the unknown signal temperature.

2.2 TEMPERATURE MEASURING RADIOMETERS

The earliest use of radiometers was the measurement of radiant temperature. These early instruments were noncoherent Dicke radiometers. At first the voltage out of the phase detector was measured directly. This voltage is at least ideally proportionate to the difference between the reference load and the antenna temperature.
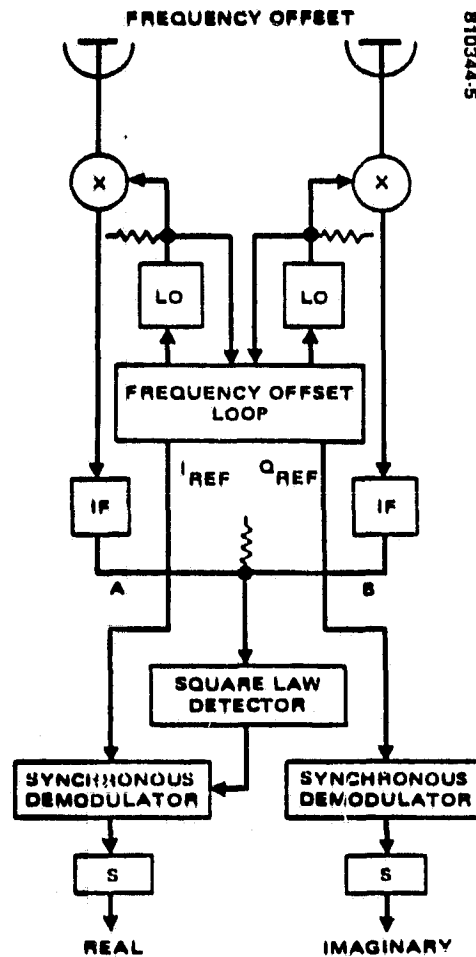
FIGURE 2-4. FULLY COHERENT
CORRELATING RADIOMETER
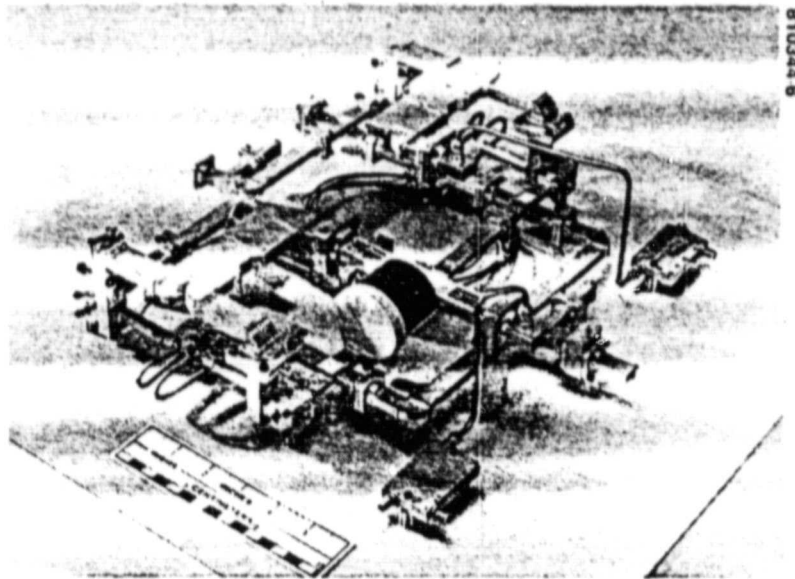
FIGURE 2-5. 1.5 cm CORRELATING RADIOMETER
RF HEAD (PHOTO 79-82146)

It was soon found that amplifiers were not stable enough to maintain an accurate ratio between antenna temperature and output voltage. The obvious answer is a servo. One approach is to servo the gain (Reference 13). The other possibility is to servo the reference load temperature to the antenna temperature by adding thermal noise power to the output from the cold load. The noise is generated by a hot standard and injected through a directional coupler (Reference 14).

The servoed attenuator scheme never went anywhere. After all, an error of 0.1 db corresponds to an error of 5°K or so. Radiometers are better at calibrating attenuators than vice versa. Neither has the servoed gain approach been too successful for reasons which are obscure.

The most successful solution has been the Hach servo loop (Reference 15). In this loop, a calibrated source, square-wave modulated, is injected into either the antenna channel or cold load channel as the case may be. Then the duty factor of the square wave is servoed to balance antenna and load channels. Initial calibration is carried out by noting the noise source duty factor using standard temperature loads. This is the system used by the radiometer in Figure 2-5.

The unknown temperature is then found by calculations using the duty factor values. The duty factor measurements are time measurements. Time measurements can be made with very great precision.

Typical results for a Hach loop radiometer are reported by Hardy et al (Reference 16) and by Love and Van Melle (Reference 17). A block diagram of the Love/Van Melle radiometer is given in Figure 2-6.

Figure 2-7 shows a measurement of the boiling point of liquid argon. It agrees to 0.1°K with the tabulated boiling point of 87.3°K given for the existing air pressure. As a matter of fact, this radiometer maintained an accuracy of 0.1°K throughout an entire flight test program without recalibration of any kind.

However, the stability of transistor amplifiers has become good enough to make temperature measurements with an unservoed radiometer, if accuracy specs are not too demanding. The radiometer is calibrated while its feed horn looks at load near 290°K, and again while the horn looks at a cold load or the sky (Reference 18). As the title indicates, the radiometer whose performance was measured was a total-power radiometer. However, this calibration technique is applicable to any radiometer including correlators.

The final type of temperature measurement of importance is the thermal-gradient mode. This is a mapping mode where the antenna scans a pencil beam fast enough to cause all the harmonics of the mapped image to occur at frequencies higher than those where gain fluctuation and 1/f noise is important.

The amplified square-law detector output is passed through a bandpass filter. This filter's upper cutoff is set by the usual considerations when designing a smoothing filter. The lower cutoff is set to reject the 1/f noise.

This simple radiometer can make excellent microwave radiometer images. Removal of the low frequencies improves the image appearance. However, loss of these frequencies prevents making absolute temperature measurements.

Figure 2-8 shows a thermal gradient radiometer designed and tested by Wiley Electronics in 1960. The antenna is a spherical reflector illuminated by three rotating line source feeds which correct for spherical aberration. This is the first mapping microwave radiometer. It made excellent images.

2.3   SPECTROMETERS

When spectrum measurement is mentioned, usually the first thought is the use of filters (Reference 19). If speed is not too important, a single scanning filter will do. In order to get more speed a bank of channelizing microwave filters can be used. This is the route taken by the writer in designing a "frequency scanning" mapping radiometer for JPL. This radiometer's amplifier traveling wave tubes (TWTs) and filter bank are shown in Figure 2-9.
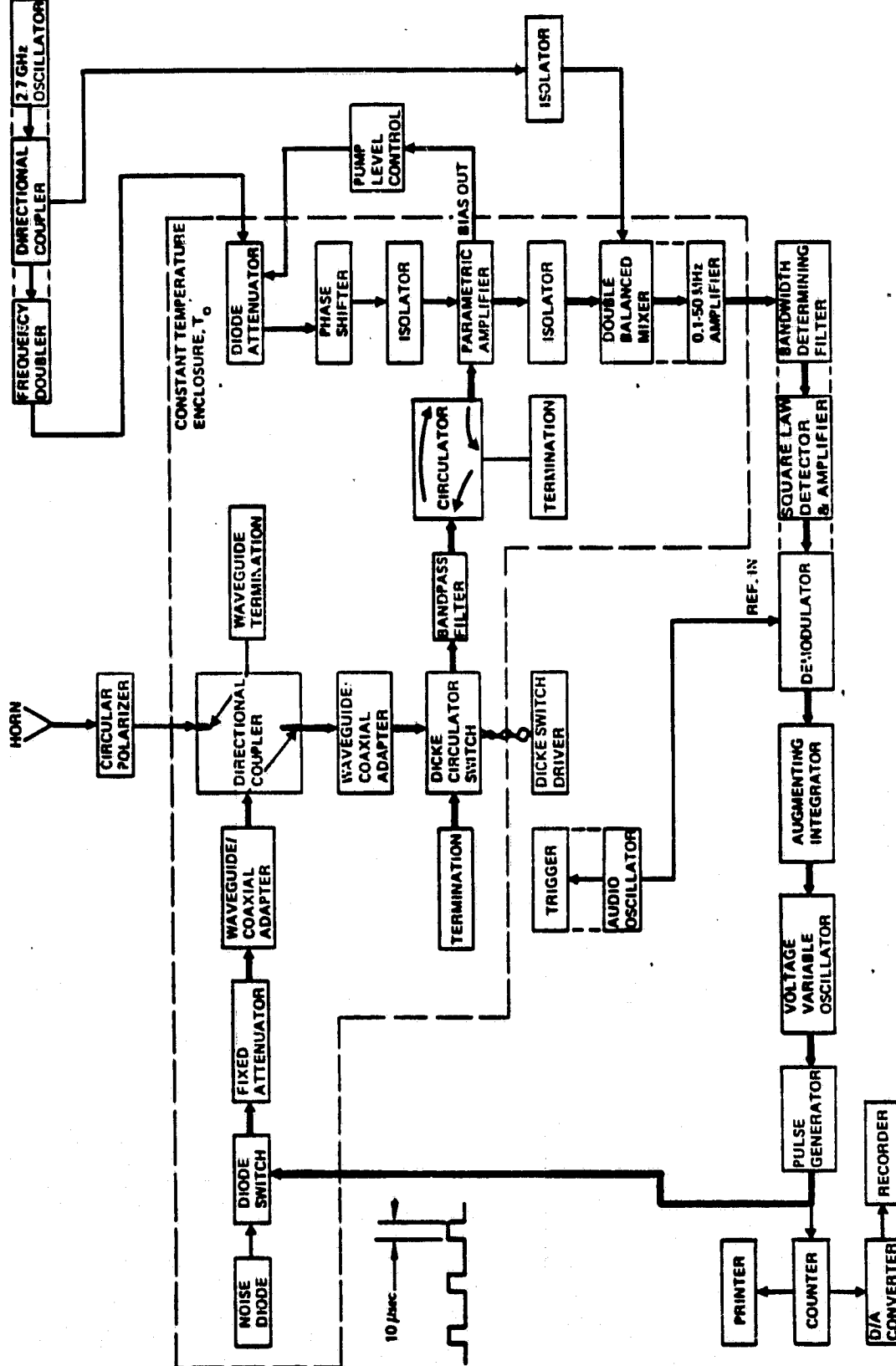
FIGURE 2-6. DETAILED BLOCK DIAGRAM OF A PRACTICAL DESIGN FOR A RADIOMETER

SYSTEM RESPONSE TIME $T = 1.0$ sec
REFERENCE TEMPERATURE $T_o = 303.2k$
ANTENNA AVERAGE TEMPERATURE $T_p = 15.0°C$
VERTICAL SCALE 1k PER INCH
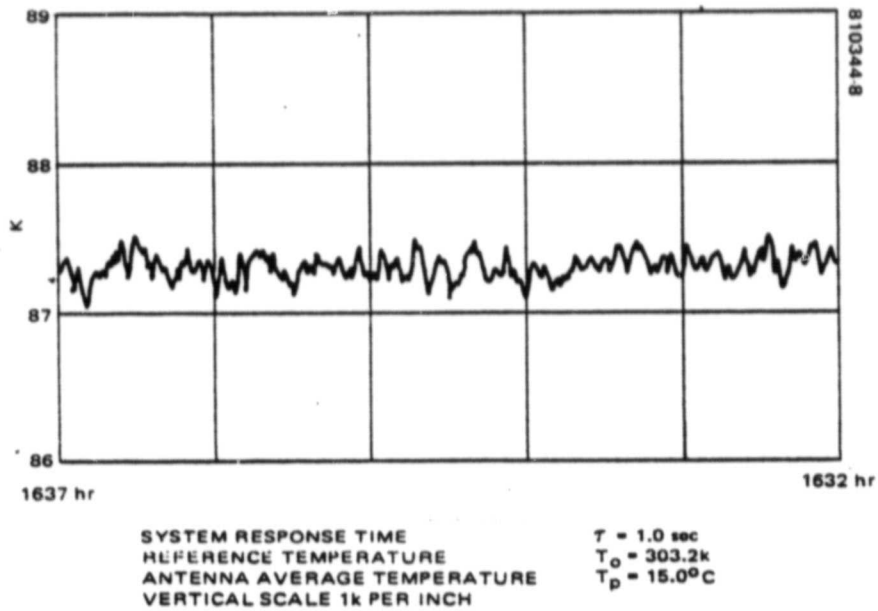
FIGURE 2-7. LIQUID ARGON TEST RECORD, 3-8-72

FIGURE 2-8. MAPPING MICROWAVE
RADIOMETER MOUNTED IN C-82
AIRCRAFT
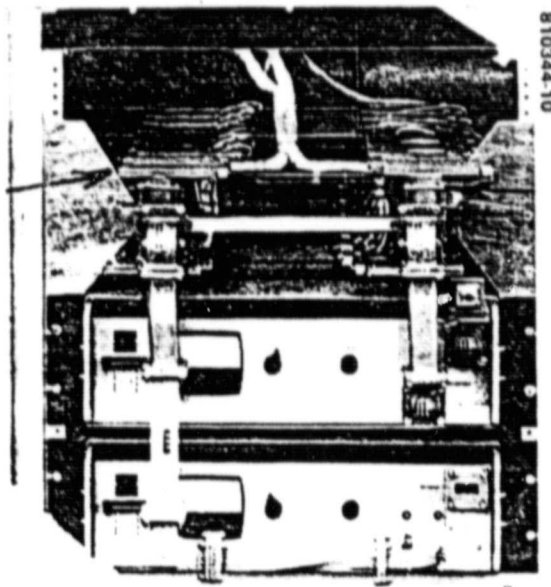
FIGURE 2-9. FIFTY-CHANNEL FILTER
BANK DRIVE BY TWO LOW NOISE TWTS

The channelizing filter banks can be seen at the top of the cabinet. Some of the disadvantages of this approach are obvious from the photograph. The filters are bulky. They are also lossy; so considerable gain between the antenna and filter bank is needed. In this case the gain was provided by two TWTs.

Another flaw appeared too late in the program to rectify. The sensitivity of the individual channels drifted with temperature changes, time, drive level, and other clauses which could not be identified.

Although the basic mapping method was proven out, this radiometer made poor images. The uneven channel reponse causes heavy streaking in the longitudinal (flight path) direction.

This radiometer was built over a decade ago. Modern components would make it possible to build a much better instrument. Feedback loops in each filter channel could eliminate uneven response. However, there is a more subtle deficiency which the competing Fourier-transform radiometers do not have.

This deficiency is the fact that the performance parameters of the spectrum estimation process are set by the properties of the filter bank and associated equipment. This means that these properties are "cast in concrete." There is no way of changing them short of rebuilding the radiometer.

The alternate spectrometer design is the Fourier-transform spectrometer. This device measures the autocorrelation function of the signal. The Fourier-transform of the autocorrelation is then found. The Wiener-Khintchine theorem states that this transform is the power-density spectrum of the signal (Reference 20).

To form the autocorrelation function the signal must be split into two components. A differential time delay between channels must be introduced. The delayed signals are then multiplied together and averaged to form $\phi(\tau)$. Figure 2-10 shows a correlator which uses a variable ultrasonic delay line. Figure 2-11 shows a correlation function measured by this radiometer. The measurement is made at various noise levels in order to measure sensitivity.

If a variable path length (trombone) is used to produce the differential delay, we have the equivalent of the scanning microwave filter. Such a device was designed and tested by Long and Butterworth in 1963 (Reference 21). They did not transform $\phi(\tau)$ to get the spectrum. Instead they drove the trombone at a steady rate. The trombone output is Doppler-shifted by the trombone movement.

A little thought will show the Doppler spectrum appearing at baseband in a replica of the microwave spectrum. Long and Butterworth therefore measured the RF spectrum by measuring the video spectrum with a low frequency spectrum analyzer. Figure 2-12 shows a block diagram of such a variable delay spectrometer. The spectrum analysis is carried out digitally in accordance with contemporary design practice.
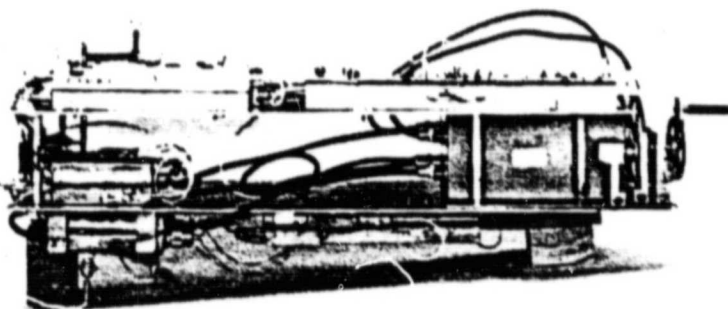
FIGURE 2-10. ELECTRONIC CORRELATOR WITH VARIABLE QUARTZ DELAY LINE TO PRODUCE DIFFERENTIAL DELAY BETWEEN TWO SIGNALS (PHOTO 700-54-3 A)
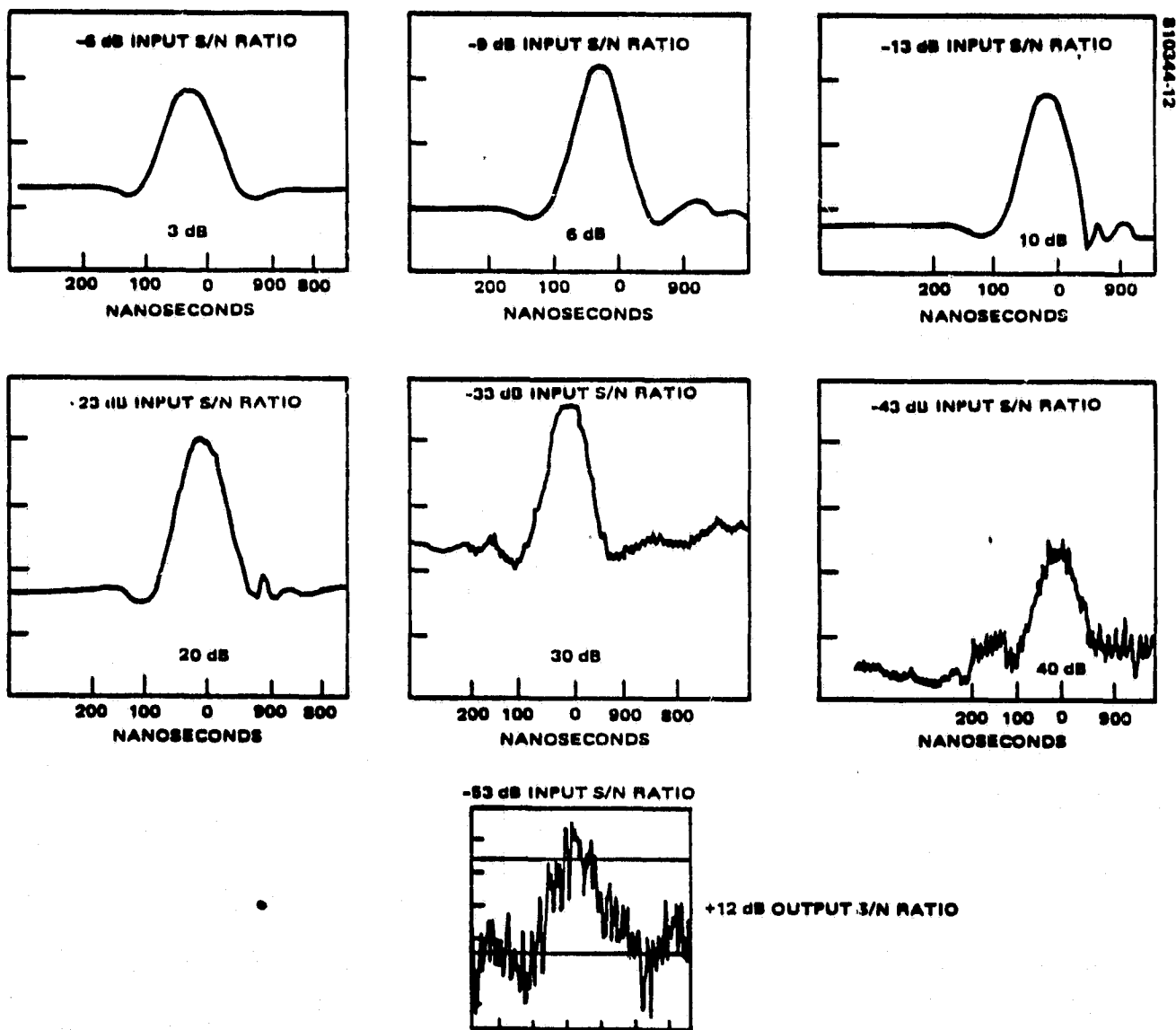
FIGURE 2-11. RECORDED CORRELATOR OUTPUT
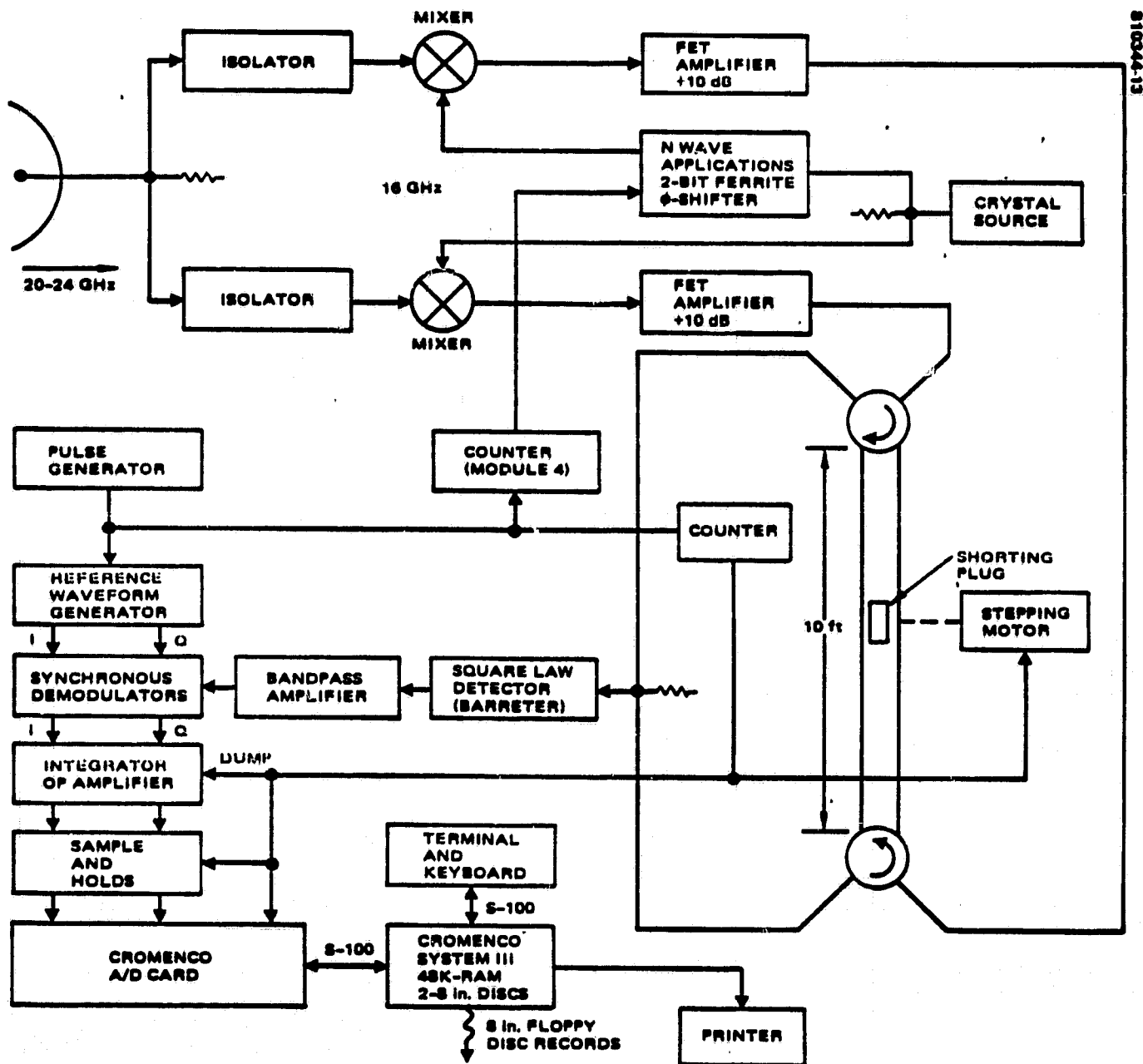
FIGURE 2-12. DEMONSTRATOR TRANSFORM SPECTROMETER

Usually the emphasis on performance dictates simultaneous measurement of all values of $\phi(\tau)$. Figure 2-13 shows such a system. The values of $\phi(\tau)$ are measured at the intermediate frequency because the correlator is very wideband (GHz). The two IF signals are fed to opposite ends of two transmission lines. At the center of the lines the delay in both channels accumulated from the point where the signals split in the hybrid, is the same. For taps on one side of the center point τ is increasingly positive with increase in distance from center. Going the other way from center τ is more and more negative.

The directional couplers tap power from both lines. The circuitry driven by each tap pair is similar to that shown in Figure 2-4 starting at points A and B. The processed tap-pair outputs are then multiplexed, digitized and fed to a computer where the Fourier transform of $\phi(\tau)$ is found.

If the phase-shifter in Figure 2-13 is correctly indexed, the outputs of the square-law detectors can be exactly in phase with the I reference. The Q-channel electronics is then not needed.

The digital values of $\phi(\tau)$ are fed to a computer where a Fourier transformation is carried out to yield the spectrum estimate.

If the bandwidth of the system is narrow enough, the signal can be converted to base-band, digitized and delayed in shift registers. The register outputs are then multiplied by the original undelayed digital signal in digital multipliers and integrated in digital accumulators to form the set of autocorrelation function values. Weinreb contrasts the analog and digital approaches in his thesis (Reference 22). His Figure 4, shown here as Figure 2-14, gives functional block diagrams for the two correlator types.

Weinreb's correlator used one-bit quantization in order to make the digital circuitry simpler. It should be remembered that, in 1963, LSI had not arrived yet. A decade later Bowers et al report on multibit correlator (Reference 23).

A tutorial paper on autocorrelation spectrometers is given by Cooper (Reference 24). This paper is completely oriented toward digital correlators. In Section 3.5.10, Cooper reviews the radio-telescope correlators operating at that time. This data is summarized in Table 2-1.

2.4 INTERFEROMETERS

The simplest type of interferometer is the adding interferometer. The output of the antennas at the end of the interferometer baseline are added in a hybrid and passed to the receiver as shown in Figure 2-11.

The interferometer pattern consists of a series of lobes. They are all positive as indicated on the pattern diagram in Figure 2-15.
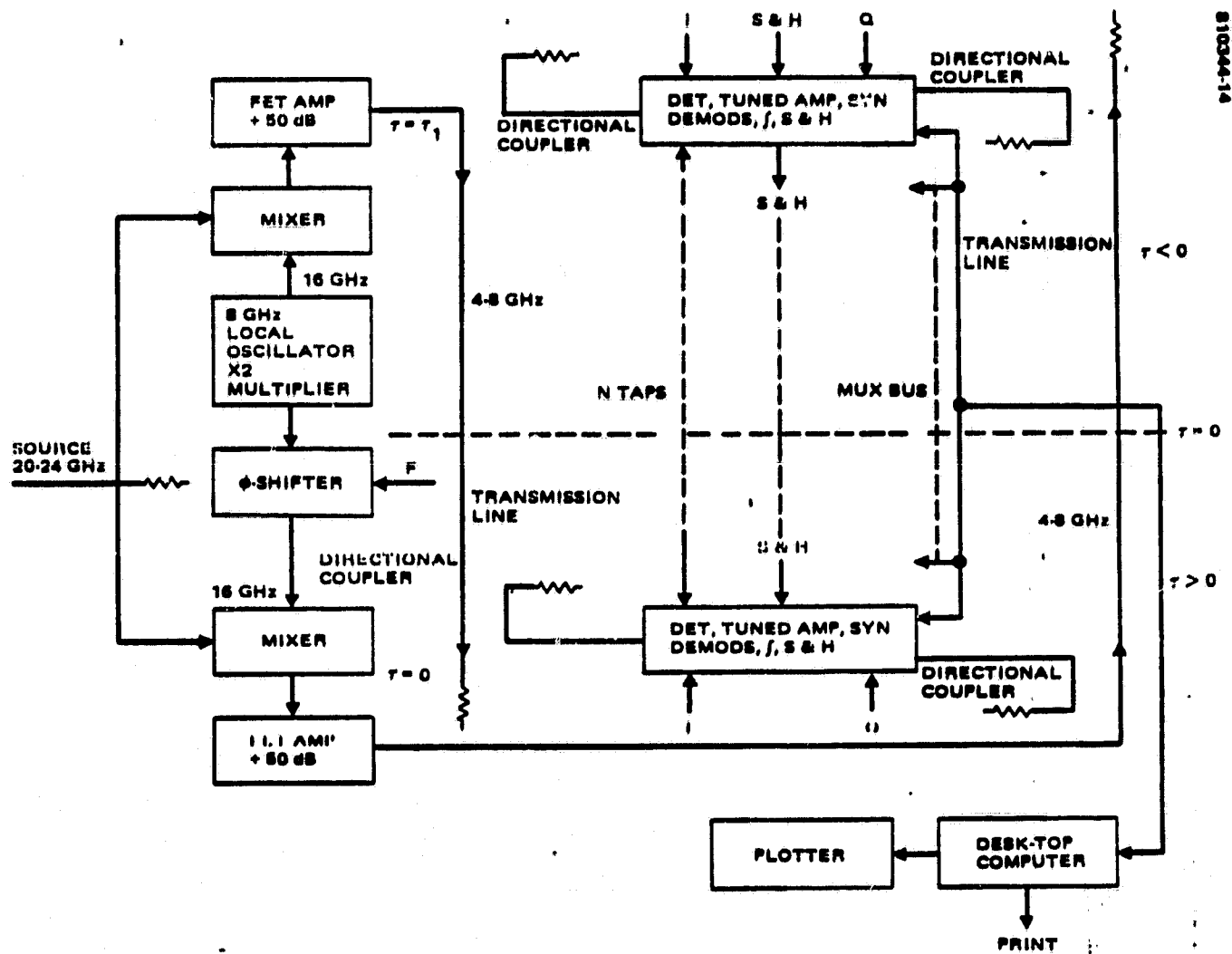
FIGURE 2-13. SIMULTANEOUS DELAY TRANSFORM SPECTROMETER
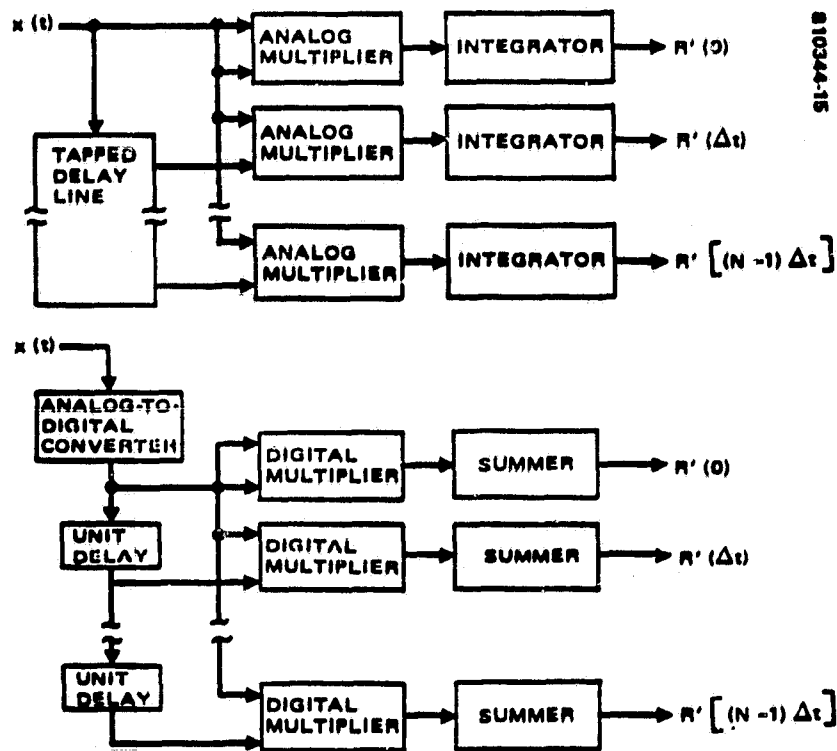
FIGURE 2-14. ANALOG AND DIGITAL AUTOCORRELATION
METHODS OF SPECTRAL MEASUREMENT

TABLE 2-1. OPERATIONAL CORRELATION SPECTROMETERS (1976)

Location	Clock Rate (MHz)	BW	Channels	Bits	Comments
NRAO, Green Bank	0.039 to 20		413		
MII, Lincoln Laboratory	50				Interleaved Sampling
U.S. National Radio Observatory	50				
Jodrell Bank Observatory		10	256		
C.S. I.R.O. Radio Physics Lab	20		1024		
Univ. of British Columbia		16	128	1	
Univ. of British Columbia		4	128	3 x 5	Dual Mode
Haystack Observatory	25	10	1024		
	50	20	512		Trimode
	100	40	256		
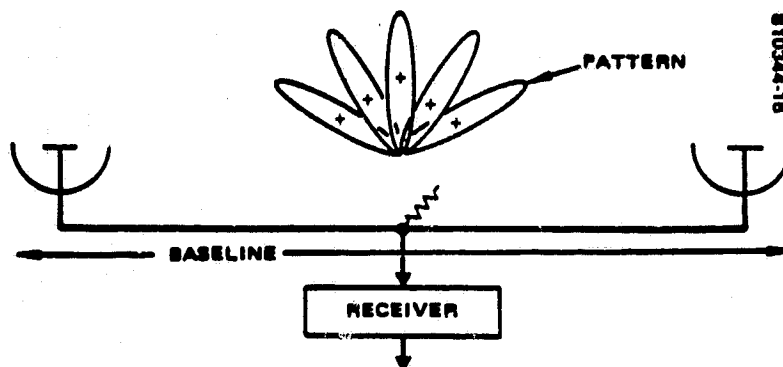
FIGURE 2-15. SIMPLE ADDING INTERFEROMETER

The D.C. component of this pattern gives no information about the scene. Its presence complicates the use of the data collected. A much better type of interferometer is the multiplying interferometer. This type can be mechanized by use of a correlator to carry out the multiplication. The phase-switched interferometer is also a multiplying interferometer. These possibilities are shown in Figure 2-16.

Note that the lobes are alternately positive and negative. There is no DC component. Figure 2-17, taken from Krause (Reference 8), shows the output of the adding and of the multiplying interferometer as a source traverses their fields of view.

a) CORRELATION FUNCTION

b) PHASE-SWITCHED
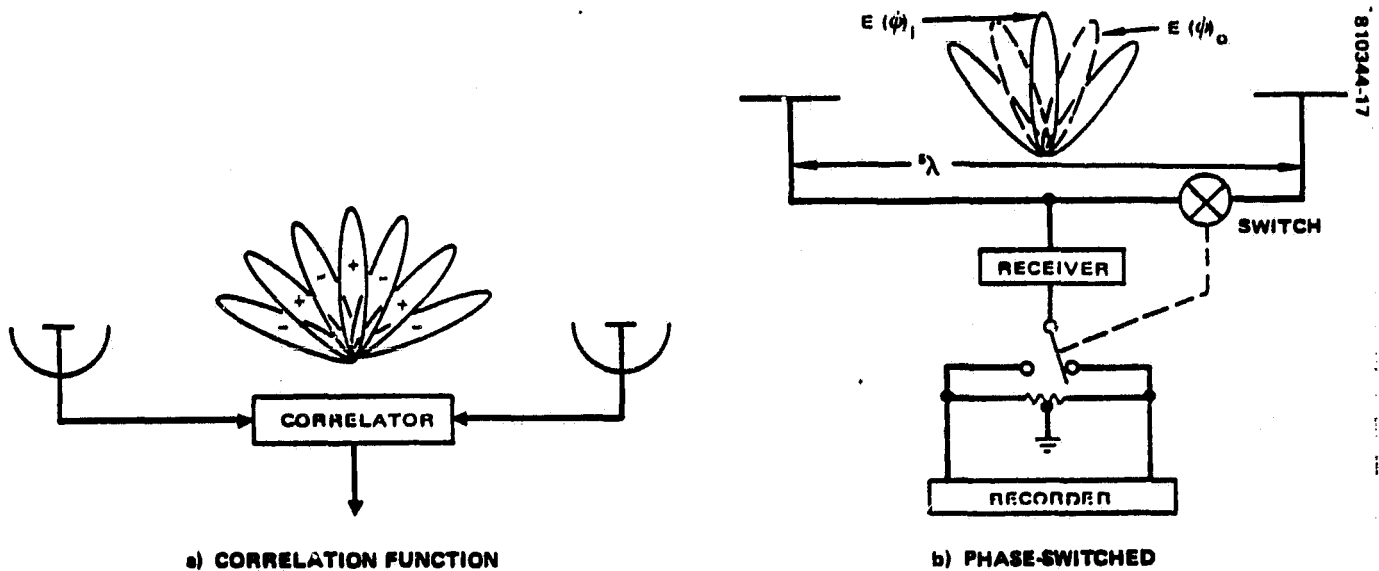
FIGURE 2-16. MULTIPLYING INTERFEROMETER

a) SIMPLE (UNSWITCHED) INTERFEROMETER OUTPUT

b) PHASE-SWITCHED INTERFEROMETER OUTPUT
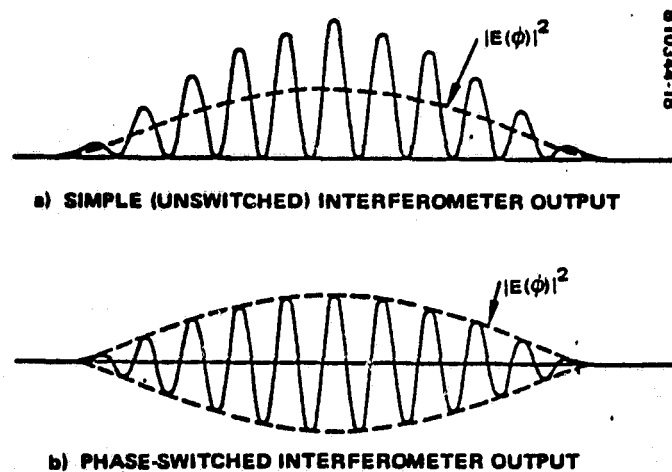
FIGURE 2-17. RESPONSES OF AN ADDING INTERFEROMETER

A significant restriction on the use of an interferometer is caused by decorrelation of the signals from the two antennas due to their differential time of arrival at these antennas. Roughly, this differential time must be less than the bandwidth of the signal.

There are several ways of easing this restriction. One popular method is used when the instantaneous required FOV is small, but moves through large angles over long periods of time. For instance, in the case of radio telescopes the FOV is often determined by fairly large parabolic antennas. The parabolas see only a small part of the sky. However, over periods of hours, this small FOV moves to another part of the sky. The solution in this case is to switch in pieces of transmission line in the path back to the correlator. By using the right line length the differential time of travel to the correlator can be made close to zero for a source at the center of the FOV.

In many cases the spectrum of the signal emerging from the interferometer antennas is desired. If the spectrum is found by passage of the signal through a bank of channelizing filters or by transform spectroscopy and cross correlation is carried out after this process, decorrelation is usually not a problem. Each spectrum channel can be made narrow enough to ensure correlation in the presence of the differential time delay between antennas signals existing as a result of other design considerations. A sketch of this scheme follows.

Obviously the arrangement in Figure 2-18 is needed for atmospheric sounding. Sounding measurements involve measurement of the spectrums of atmospheric absorption lines. A radiometric version of frequency scanning will be discussed later in this report. It also uses a multichannel system configuration. The types of correlation functions which can be generated is discussed in E. Tensor Synthesis.
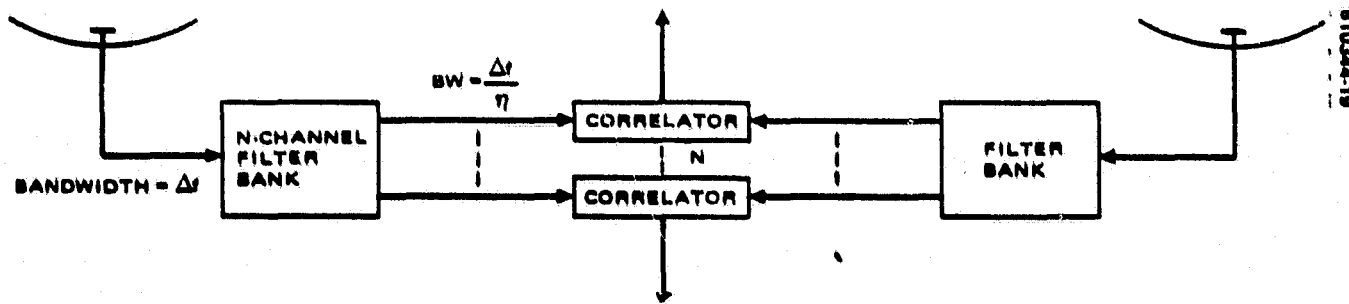
FIGURE 2-18. MULTICHANNEL MULTIPLYING INTERFEROMETER

For the sake of completeness intensity interferometry should be mentioned. This technique detects the signals in the two channels first. Then the envelopes of the two video signals are correlated. It can be shown that the correlation coefficient of the squares of the moduli of the antenna voltages (detected video) is equal to the square of the correlation of the incident radiation (Reference 25). As one would suppose, this eliminates the affects due to the RF structure of the radiation. Therefore, correlation can be obtained over much greater baseline lengths. Also, no phase information need be transmitted to the signal receiver modules at the ends of the baseline. The detected signals, therefore, can be recorded on tape and correlated at some site remote from the antenna sites.

This technique is very-long baseline interferometry (VLBI).

However, removal of the signal RF structure from the correlation process exacts a heavy penalty. The sensitivity is considerably worse than RF cor-relators (Reference 26). Even after reading some review papers, it is not clear to the writer that VLBI can produce images (Reference 27), rather than measurements of point sources. Therefore, VLBI will not be considered further in this study.

2.5 POLARIMETERS

In many cases, the polarization state of received radiation is desired as well as the quantities already discussed. Polarization can be described by several methods. One description is the Stokes vector which is favored in optics.

If a_1 and a_2 are the amplitudes of the two linearly polarized orthogonal components of radiation, then the Stokes vector is

$$
S = \begin{vmatrix} \langle a_1^2 + a_2^2 \rangle \\ \langle a_1^2 \quad a_2^2 \rangle \\ \langle 2a_1 a_2 \cos \gamma \rangle \\ \langle 2a_1 a_2 \sin \gamma \rangle \end{vmatrix} = \begin{vmatrix} I \\ M \\ C \\ S \end{vmatrix} \qquad (I\text{-}2\text{-}1)
$$

where γ is the phase angle between a_1 and a_2. This assumes that the radiation is quasi-monochromatic. Then γ can be defined and measured even for noise-like signals.

Consider three pairs of antennas: let two antennas be linearly polarized at 0° and 90°; linearly polarized at 45° and 135°; finally, two antennas right and left circularly polarized.

The powers available from these antennas are denoted by P_0, P_{90}, P_{45}, P_{135}, P_R, and P_L, respectively. In optics, these same quantities are self measured by dichroic filters, which extract various linear polarization components from a beam of light. They are preceded by a retardation plate. Then I(0, c) denotes the intensity of light vibrations in the direction making an angle with coordinate axis OX, when the y-component is subjected to a retardation ε with respect to the x-component.

Then the components of the Stokes vector are (Reference 28)

$$I = I(0°, 0) + I(90°, 0) = P_0 + P_{90}$$

$$M = I(0°, 0) - I(90°, 0) = P_0 - P_{90}$$

$$\text{(I-2-2)}$$

$$C = I(45°, 0) - I(135°, 0) = P_{45} - P_{135}$$

$$S = I(45°, \pi/2) - I(135°, \pi/2) = P_R - P_L$$

The Stokes parameters have been satisfactory in optics. However, they produce complications when used in microwave interferometry. This is due to the fact that all conceivable feed arrangements for a two-element interferometer response to a maximum of two out of four Stokes parameters (Reference 29). (Ko's letter says "minimum" instead of "maximum." This must be a typo. Equations (I-2-2) show that four antennas per baseline and are needed.)

The coherency matrix is better adapted for use with microwave radiometers. It needs only two antennas per station for measurement of two orthogonal polarizations. If these polarizations are linear, the components of the matrix ρ can be measured by an arrangement such as is shown in Figure 2-19 below (Reference 30). The power output from dipole A is ρ_{xx}. The output from dipole B is ρ_{yy}. The real part of ρ_{xy} is obtained by cross correlating the voltages from A and B. The imaginary part of ρ_{xy} is obtained by delaying the signal from A by $\lambda/4$ and then cross correlating.
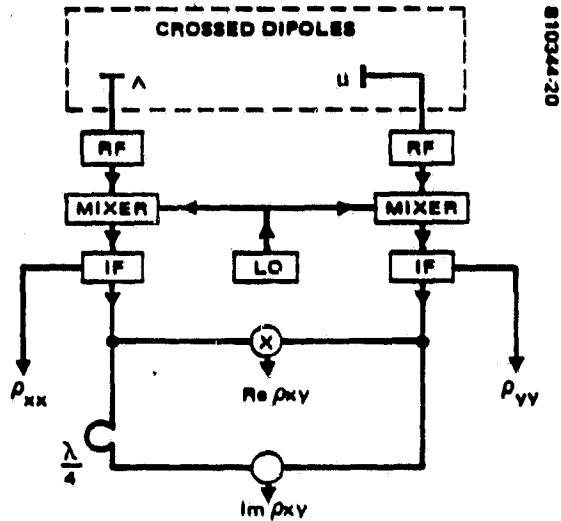
FIGURE 2-19. CORRELATION POLARIMETER
WHICH SIMULTANEOUSLY MEASURES THE
INDEPENDENT PARAMETERS OF THE COHERENT
MATRIX (THE FEED CONSISTS OF A PAIR OF
CROSSED DIPOLES)

We see then that the coherency matrix has the form

$$
\begin{bmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{bmatrix} = \begin{bmatrix} E_2 E_2^* & E_x E_y^* \\ E_y E_x^* & E_y E_y^* \end{bmatrix} = \begin{bmatrix} \langle a_1^2 \rangle & \langle a_1 a_2 \, e^{i(\phi_1 - \phi_2)} \rangle \\ \langle a_1 a_2 \, e^{-i(\phi_1 - \phi_2)} \rangle & \langle a_2^2 \rangle \end{bmatrix} \qquad (I-2-3)
$$

where

$$
E_x(t) = a_1(t) \, e^{i[\phi_1(t) - 2\pi \overline{\nu} t]}
$$

and

$$
E(t) = a_2(t) \, e^{i[\phi_2(t) - 2\pi \nu t]}
$$

The trace of J is $J_{xx} + J_{yy} = [E_x E_x^*] + [E_x E_x^*]$. The nondiagonal elements are conjugates of each other, that is $J_{xy} = J_{yx}^*$. Therefore, J is Hermitian.

The interrelation between the Stokes parameters and coherency matrix elements are given below.

$$
\begin{cases}
I = J_{xx} + J_{yy} \\
M = J_{xx} - J_{yy} \\
C = J_{xy} + J_{yx} \\
S = i(J_{yx} - J_{xy})
\end{cases}
\qquad
\begin{cases}
J_{xx} = 1/2(I + M) \\
J_{yy} = 1/2(I - M) \\
J_{xy} = 1/2(c + is) \\
J_{yx} = 1/2(c - is)
\end{cases}
\qquad (I\text{-}2\text{-}4)
$$

Often we may wish to measure the polarization state of each image pixel. If the image is generated by aperture synthesis, the basic interferometer configuration due to Ko (Reference 30) follows on the next page.

Both Figure 2-19 and Figure 2-20 show homodyne correlating radiometers. In a practical system it is better to use a fully-coherent radiometer. One advantage will now become apparent. It was mentioned in the spectrometer section that the signal whose spectrum is measured could be contained solely in the in-phase signal channel by proper adjustment. This means that a second signal can be passed through the quadrature channel of the same instrument. The natural signal to inject into the two channels are the pair of orthogonal polarizations. In most remote sensing applications the radiant brightness of both polarizations are required. So dual use of the same radiometer considerably simplifies the sensor.

Passage of the dual signals through mixers is the most subtle operation required in order to mechanize the dual radiometer. Figure 2-21 shows one embodiment of such a mixer.

The LO signals in both channels (separated in frequency by the off-set frequency) are injected into the mixer input ports by means of a coupled-resonator filter as shown. The normal operation of the mixer then delivers E_A to one IF port and E_B to the other port.
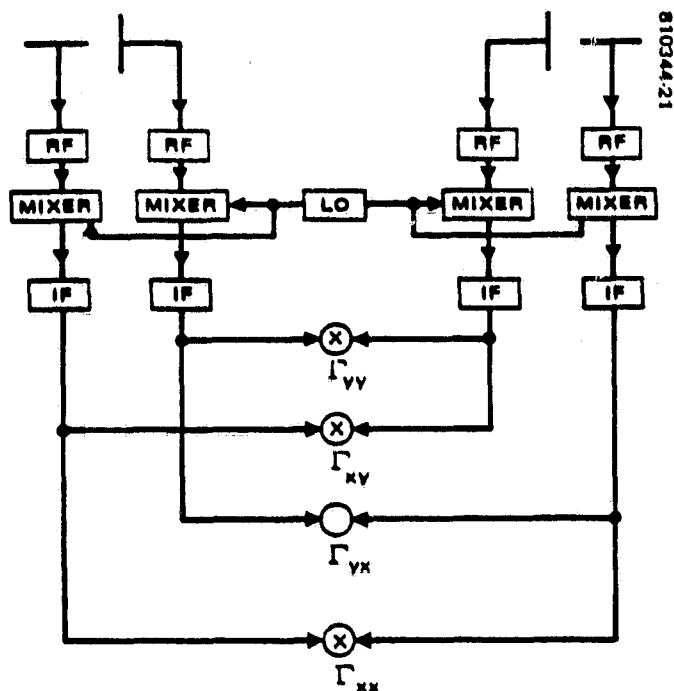
FIGURE 2-20. INTEFEROMETRIC POLARIMETER WHICH SIMULTANEOUSLY MEASURES THE FOUR POLARIZATION COHERENCE FUNCTIONS FOR THE FOURIER SYNTHESIS OF THE POLARIZATION AND BRIGHTNESS DISTRIBUTION
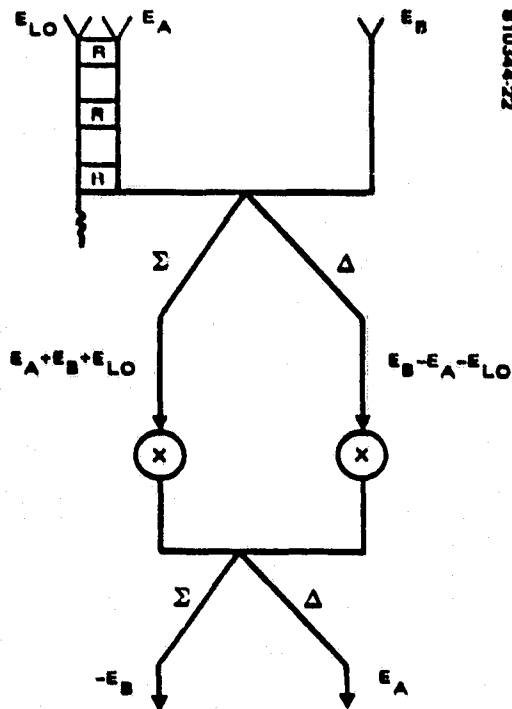
FIGURE 2-21. DUAL SIGNAL BALANCED MIXER

Figure 2-22 shows the total dual-polarization radiometer. Note that E_x is split into two inputs to the correlator channels by means of an in-phase hybrid. E_y is split into its channel input signals by means of a quadrature hybrid. Therefore, we have E_x and E_y in the left channel the right channel has E_x and E_y. The sum and difference of these signals are formed in both channels by the hybrids which drive the mixer diodes. In the left channel $E_x + jE_y$ goes to one diode and $E_x - jE_y$ goes to the other diode. The output hybrid then forms E_x and jE_y at its sum and difference ports after the signals are reduced to the intermediate frequency. Likewise E_x and E_y are formed at separate IF mixer ports in the right channel.

The correlators are mechanized as shown in Figure 2-23.

From Figure 2-22 we see that $E_x E_x^*$, $E_y E_y^*$, $(E_y E_x^*)_{imag}$, and $(E_y E_x^*)_{real}$ are formed. These are the elements of the coherence matrix.

The servos indicated in the block diagram are Hach servo loops which have been discussed before. A functional block diagram is given in Figure 2-24.

The square wave noise bursts are injected into correlator arms with correct phase to act like E_x, E_y and jE_y. The VCO seeks a frequency where the duty cycles of the noise diode outputs balance the correlator outputs to zero. These are Clapp calibrators. They will be discussed in more detail in a later section.

2.6   TENSOR SYNTHESIS

Ko (References 29, 30) calls the simultaneous measurement of the many coherence functions associated with the incoming radiation simultaneously "tensor synthesis." We will not treat the subject as formally as he does in Reference (30).

It should be noted that the most general space-time coherence function, $\Gamma(x_1, x_2, \tau)$, involves cross correlating signals at locations x_1 and x_2 with differential time delay τ. Then $\Gamma(x_1, x_1, \tau)$ is a temporal (longitudinal) coherence function. Its Fourier transform is the radiation power spectrum. We also have a lateral coherence function $\Gamma(x_1, x_2, 0)$. Its transform is the brightness distribution of the radiation source according to the van Cittert-Zernike theorem.

Returning to the general coherence function we see that a generalization of these transforms implies that $\Gamma(x_1, x_2, \tau)$ has associated with it three other functions. Namely, its "spatial" Fourier transform, $\tilde{\Gamma}(\mu_1, \mu_2, \tau)$; its "temporal" Fourier transform $\Gamma(x_1, x_2, \nu)$; and its total Fourier transform $\Gamma(\mu_1, \mu_2, \nu)$.
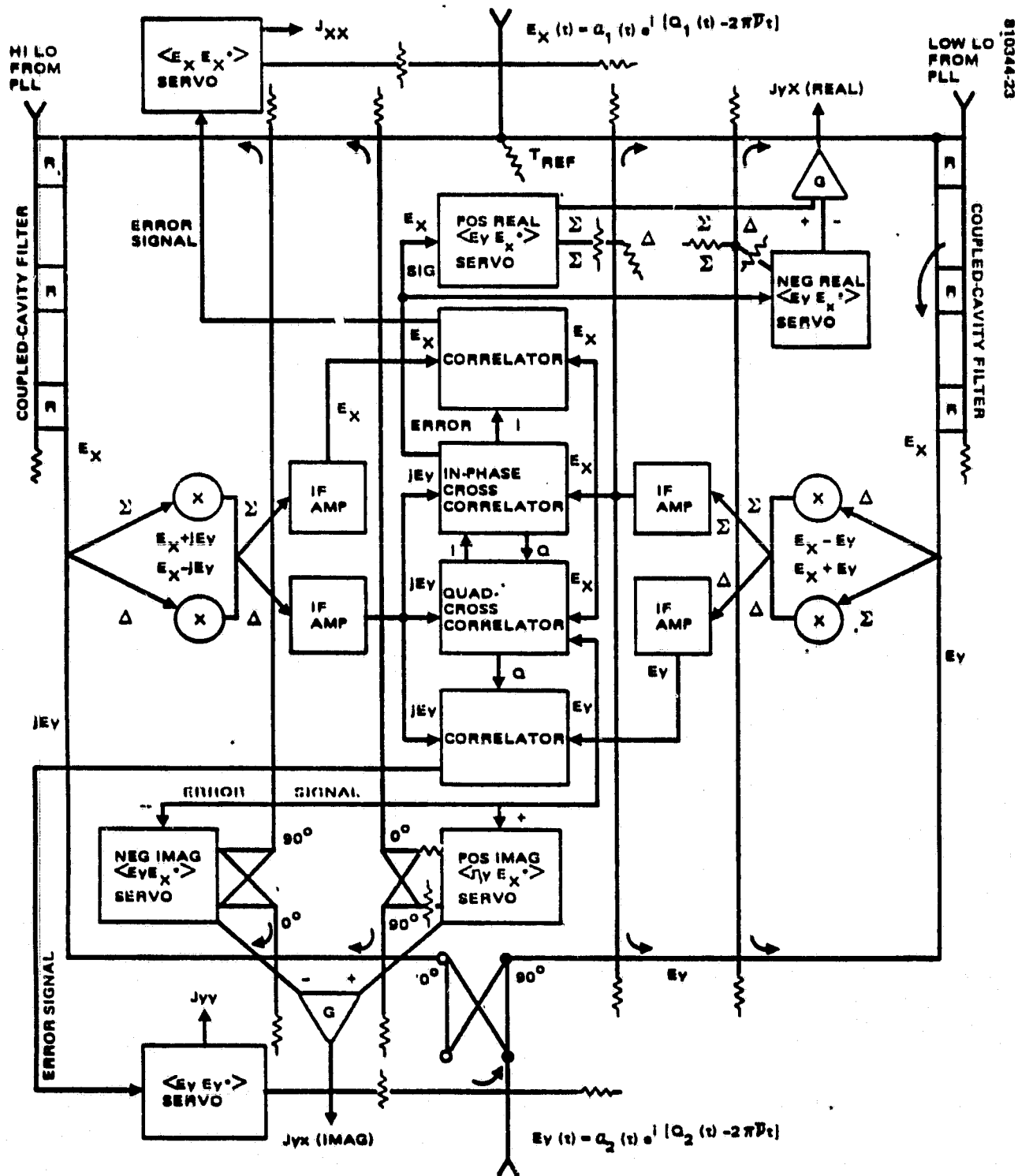
FIGURE 2-22. CORRAD CONFIGURATION FOR COHERENT MATRIX MEASUREMENT

E_A E_B

POWER
COMBINER

SQUARE
LAW
DETECTOR

I OR Q REFERENCE
FROM LO OFFSET SERVO LOOP

BANDPASS
FILTER
AT LO OFFSET
FREQUENCY

SYNCHRONOUS
DEMODULATOR
(PHASE
DETECTOR)

INTEGRATOR
FILTER

$\overline{E_A E_B}$
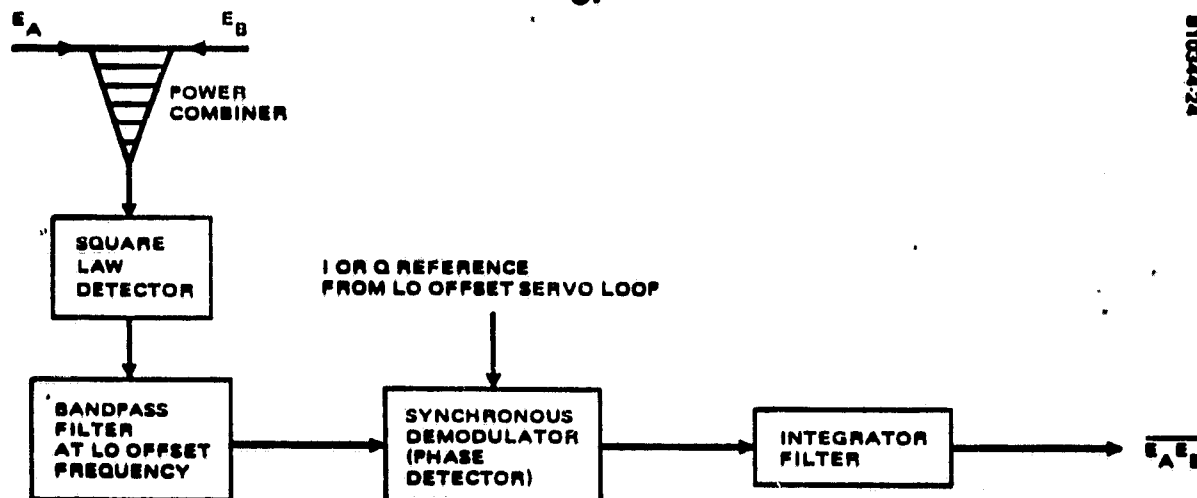
FIGURE 2-23. CORRELATOR CONFIGURATION

810344-24

EVENTS-
PERUNIT-
TIME CKT

VALUE OF
CORRELATION

810344-25

ERROR
SIGNAL

VCO

ZERO-
CROSSING
PULSE
GEN

ONE-
SHOT
MULTI-
VIBRATOR

NOISE
DIODE

OUT
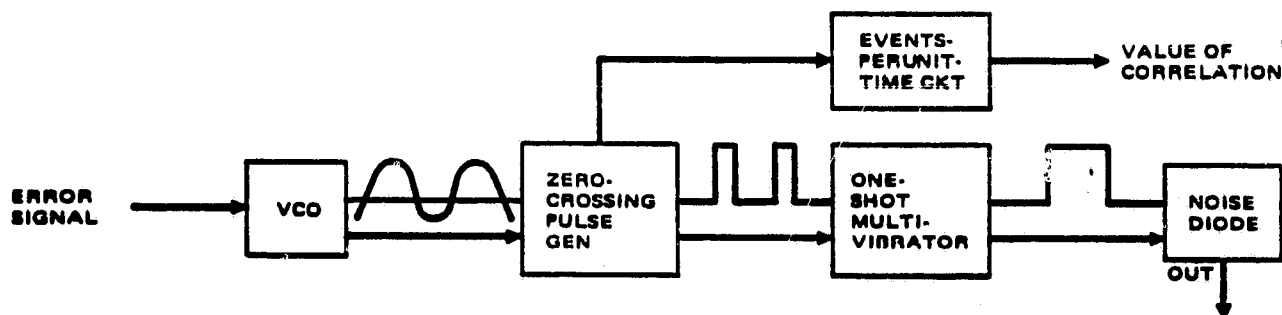
FIGURE 2-24. HACH SERVO LOOP

The question arises then as to how we can generate the coherent functions involved and their transforms. In order to do this assume that antennas are located at two positions. Let us label the various variables involved as follows:

i, j   Polarization components

k,   Time

m, n   Position (m = y position change, n = z position change)

p, q   Frequency

y, z   Spatial frequency components in y and z directions

Let the correlators at positions m and n consist of pairs of transmission
lines, to provide delay, together with output circuitry as shown in Figure 2-25.
Let these correlators have polarization inputs E_i and E_j. The cross correlation outputs are $\Gamma_{ij}^m(\tau k)$ and $\Gamma_{ij}^n(\tau \ell)$. The Γ^i subscripts indicate which element of the polarization coherency matrix is represented by the Γ. The superscripts indicate the locations of the interferometer baseline. The τ subscripts indicate the set of differential delays available. The temporal transforms of
the two correlator outputs are $\hat{\Gamma}_{ij}^m(\nu_p)$ and $\hat{\Gamma}_{ij}^n(\nu_\phi)$.

If movement of the interferometer baseline generates a series of
y-position changes, x_m, and a series of z-position changes normal to y displacements, we have a set of correlation measurements at x_m, x_n. If correlator outputs with a delay of $\tau = \tau_k = \tau_\ell$ are correlated, we get $\Gamma(x_m, n, \tau)$. Its
temporal transform is $\hat{\Gamma}_{ij}(x_m, x_n)$. Its spatial transform is
Finally we can transform $\Gamma_{ij}^m(\tau_k)$ and $\Gamma_{ij}^n(\tau_\ell)$ to form $\hat{\Gamma}_{ij}^m(\nu_p)$ and $\hat{\Gamma}_{ij}^n(\nu_\phi)$.
If a set of displacements x_n and x_m have been generated and Γ^m and Γ measured
for these displacements, we can set $\nu_p = \nu_\phi = \nu$. Then the two-dimensional
spatial transform gives $\hat{\Gamma}_{ij}(x_m, x_n, \nu)$. Its transform is $\Gamma_{ij}(\mu_y, \mu_z, \nu)$.

A block diagram of these manipulations is shown in Figure 2-25.
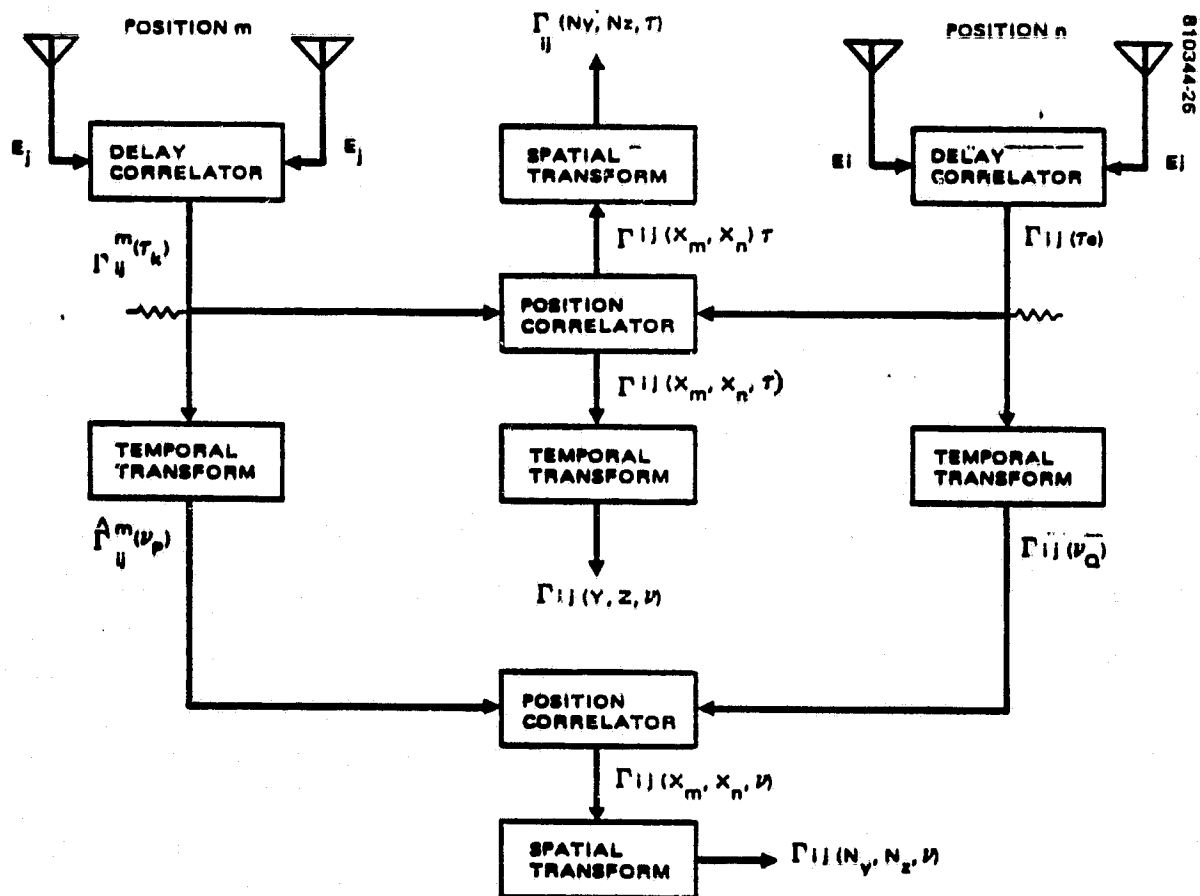
FIGURE 2-25. GENERATION OF COHERENCE FUNCTIONS ASSOCIATED WITH (x_1, x_2, τ)

The correlation formed at the bottom of the diagram is very important. The bandwidth of any frequency channel is $\Delta B = \nu_{p+1} - \nu_p = \nu_{q+1} - \nu_q$. This bandwidth can be made small enough to ensure correlation no matter what how long the interferometer baseline is. This is important if we wish to build an aperture-synthesis system capable of atmospheric sounding. The required system bandwidth may have to include the entire O_2 absorption spectrum around 60 GHz. This is a bandwidth of 10 GHz or more.

A method of frequency scanning radiometers will be introduced in a later section. It can provide resolution and mapping normal to an aperture-synthesis axis. This method requires 10% to 20% bandwidths. As in the sounding case, this large bandwidth would destroy correlation in the aperture synthesis process, if spectrum analysis were not carried on first.

Measurement of the coherency matrix has an important function even when the polarization data is not needed for remote sensing purposes. Often the vertical and horizontal linear polarization components of the incoming radiation are desired. Frequently conical scan is used in order to ensure that the beam makes a constant angle with the terrain for all scan angles. Such a scan can be produced by a rotating reflector as shown in Figure 2-26.

If the reflector assembly rotates but the feed horn does not, the polarization components received by the feed horn change with scan angle. For instance in Figure 2-26 above, the electric field vector is shown in the plane of the paper for the reflector position shown. This is "vertical" polarization. However, when the reflector rotates 90° and the horn does not, the E vector will be normal to the ray paths shown on the drawings. It then is horizontally polarized.
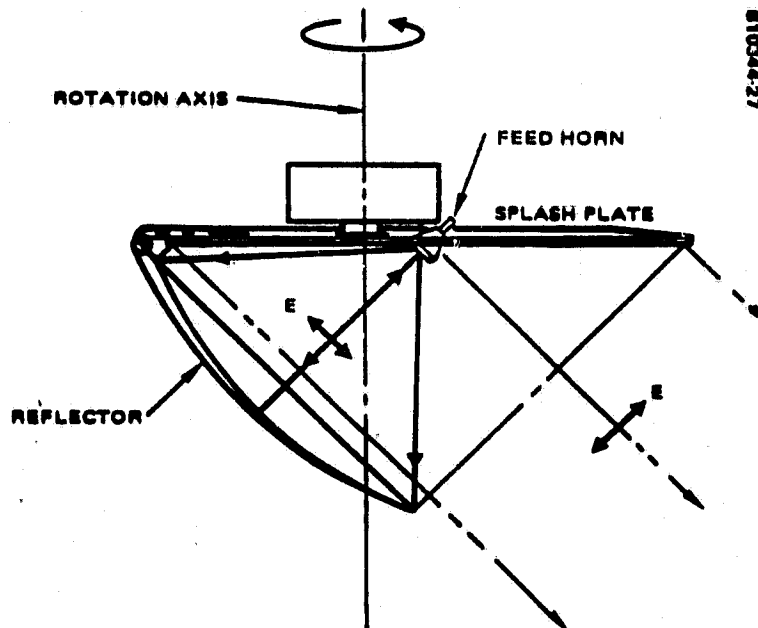
FIGURE 2-26. CONICAL SCAN, REFLECTOR ANTENNA

If the system is relatively narrow band, a rotary joint can be used between the horn and the rest of the microwave plumbing. Then it can rotate while the rest of the radiometer is stationary. However, in many systems the feed horn is a scalar horn, operating in many frequency bands, spread over a very great frequency range. There are no rotary joints which can pass such a set of operating bands.

However, if the radiometers monitoring these bands are designed to measure the coherency matrix, conversion of polarization components to correct coordinate system can be carried out by the data processing computer in the process of image restoration. This is done by multiplying the original Stokes vector [IMCS] by a rotator matrix T(C) to produce the required polarization components. This relation is

$$
\begin{bmatrix} I' \\ M' \\ C' \\ S' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\delta & \sin 2\delta & 0 \\ 0 & -\sin 2\delta & \cos 2\delta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} I \\ M \\ C \\ S \end{bmatrix}
\qquad (I-2-5)
$$

Thus we are able to have the simplicity offered by only rotating the reflector by computing the required components in the radiometer data processing section.

3. CORRELATING RADIOMETER TEMPERATURE SENSITIVITY

3.1 RADIOMETER FIGURE OF MERIT

The subject of radiometer temperature sensitivity will be touched on very lightly. For some reason sensitivity has attracted more than its share of the attention in microwave radiometer studies (References 26, 31, 32, 33, 34). In reality, statistical bias errors are more numerous, harder to control and usually larger than the temperature resolution error. Absolute temperature calibration errors are also liable to be greater than the resolution error. The temperature resolution of a radiometer is given in the two forms below:

$$\Delta T = \frac{K\, T_{SYS}}{\sqrt{B\tau}} = \sqrt{2}\, K\, T_{SYS} \sqrt{\frac{b}{B}} \qquad \qquad (I-3-1)$$

where T_{SYS} = system noise temperature, B is bandwidth before detection, b is the final bandwidth after detection, τ is the integration or smoothing time after detection, and K is the radiometric figure-of-merit. The system noise temperature is usually cast in the form

$$T_{SYS} = T_{REC} + T_{ANT} \qquad \qquad (I-3-2)$$

T_{REC} is the effective noise temperature at the end of the feedline from the receiver which connects to the antenna. T_{ANT} is the antenna background temperature. This temperature is the scene radiant brightness, weighted by the antenna cross section, integrated over a sphere surrounding the antenna and divided by 4π. The result of the averaging operation is then corrected to account for antenna ohmic losses. So

$$T_{ANT} = (L_{ANT} - 1)\, T_o + \frac{L_{ANT}}{4\pi} \int_{sphere} \sigma(0, \phi)\, B(0, \phi)\, d\Omega \qquad (I-3-3)$$

where L_{ANT} is the ohmic loss in the antenna, T_0 is ambient temperature, $B(\theta, \phi)$ is scene radiant brightness and $\sigma(\theta, \phi)$ is antenna cross-section in meter squared per steradian.

The receiver temperature is

$$T_{REC} = (L_F - 1) T_0 + L_F T_{RF} + \frac{L_F T_M}{G_{RF}} \qquad (I\text{-}3\text{-}4)$$

where L_F = feed-line loss, T_{RF} = noise temperature of preamplifier, G_{RF} = pre-amp gain, T_M = mixer-and-IF noise temperature. With sufficient preamp gain the third term becomes negligible.

When there is no preamp, the mixer and IF noise temperature dominate. So

$$T_{REC} = T_M = (F_n - 1) T_0 \qquad (I\text{-}3\text{-}5)$$

where F_n is the noise figure of the mixer/IF-amplifier.

If there is no mixer, we have a tuned-radio-frequency (TRF) receiver. Then T_M and F_n in (I-3-5) refer to the main amplifier. In superhet radiometers without image rejection both RF sidebands are received and folded on top of each other in the RF amplifier. Then the double sideband noise figure, F_{DSB}, should be used instead of the single sideband noise figure, F_{SSB}. It is often stated that $F_{DSB} = F_{SSB}/2$. However, the correct relaction is (Reference 35)

$$F_{DSB} = \frac{F_{SSB}}{2} + 1 \qquad (I\text{-}3\text{-}6)$$

When $F_{SSB} \gg 1$, the approximation is usable. However, when modern low-noise receivers are used, the approximation penalizes single-sideband receivers.

The low-frequency bandwidth, b, in (I-3-1) is the noise bandwidth of the filter/integrator. A rectangular filter which gives the same noise output power as the actual filter, and has a gain equal to the peak gain of the filter, G_{max}, has the noise bandwidth, B_n. That is

$$B_n = \frac{\int_0^\infty G(f)\, df}{G_{MAX}} = \frac{\text{Area Under Filter Gain Curve}}{\text{Maximum Power Gain of Filter}} \qquad (I\text{-}3\text{-}7)$$

Table 3-1 relates the 3 dB bandwidths to noise bandwidths for output low pass filters.

Often one sees the noise bandwidth given as the bandwidth to use for the high-frequency receiver bandwidth. This is not correct. The correct bandwidth to use is called by Evans and McLeish the radiometer reception bandwidth (Reference 31). This bandwidth is defined as

$$B_{REC} = \frac{\left\{\int_0^\infty G(f)\ df\right\}^2}{\int_0^\infty G^2(f)\ df} \tag{I-3-8}$$

TABLE 3-1. PROPERTIES OF LOW-PASS FILTERS

Filter Type	$B_{3\ dB}$	B_{NOISE}	$\dfrac{B_{NOISE}}{B_{3\ dB}}$	τ
Ideal Low-Pass	B	B		$\dfrac{1}{2B}$
RC Filters with Time Constant of τ_{AC}				
Single RC	$\dfrac{1}{2\pi RC}$	$\dfrac{1}{4RC}$	1.57	$2\tau_{AC}$
Dual RC	$\dfrac{0.643}{2\pi RC}$	$\dfrac{1}{8RC}$	1.22	$4\tau_{RC}$
Triple RC	$\dfrac{0.510}{2\pi RC}$	$\dfrac{1}{32R6}$	1.16	$6\tau_{RC}$
Gaussian (∞ RC)			1.07	$\dfrac{1}{2.12\ B_{3\ dB}}$
Box Car Filter (Ideal Integrator)	$\dfrac{0.443}{\tau}$	$\dfrac{1}{2\pi}$	1.13	
5-Pole Butterworth			1.01	
5-Pole Bessel			1.04	

Bendat and Piersol (Reference 36) call the same quantity the equivalent statistical b andwidth. The most sophisticated treatment is given by Bracewell (Reference 37), who uses the autocorrelation bandwidth. Table 3-2 lists reception bandwidths defined by Evans and McLeish.

The most detailed treatment of the sensitivities of correlating radiometers is given by Faris (Reference 38). For the homodyne type of correlator, where the correlation output appears at DC, we find K to be $\sqrt{2}$. This is given in his Eq (Reference 22). This result is reasonable. K for a single-channel, total-power radiometer is unity. The correlator has two channels. Therefore, the noise power in the output is doubled. The measured temperature is proportional to the square root of the power after the signal passes through a square-law detector or multiplier in a autocorrelator. Therefore, the correlator K will be $\sqrt{2}$.

It is unfortunate that Faris did not go on to find the value of K for a fully coherent correlating radiometer. The reader will recall that this radiometer type first heterodynes the incoming signal down to an intermediate frequency. The IF signals are then summed in a power combiner and detected by a square-law detector. The DC terms out of the detector are rejected by a bandpass filter. Due to a frequency offset introduced between the IF channels, the signal appears on a carrier and is passed by the bandpass filter. The filter output is then synchronously demodulated in two phase detectors. The phase references are in-phase and in quadrature with the IF frequency offset.

TABLE 3-2. RECEPTION BANDWIDTHS OF BANDPASS FILTERS

Filter Type	$B_{RECEPTION}/B_{3\ dB}$	$B_{NOISE}/B_{3\ dB}$
Cascaded SYN tuned CKTS	3.14	1.57
Number of Stages		
1	3.14	1.57
2	1.96	1.22
3	1.76	1.15
4	1.62	1.13
(∞ Gaussian)	1.50	1.07
Rectangular	1.00	1.00

It should be obvious that synchronous detection rejects the system noise in quadrature with the reference signal. Furthermore, when the noise originates in two receivers it is evident that the noise power in phase and in quadrature with the reference will be equal. Therefore, the in-phase and quadrature baseband signals resulting from phase detection will have half the noise power which emerges from the bandpass filter. The corresponding value of K, is unity.

However, those who want a more rigorous proof of these facts will find them given by Lindsey (Reference 39).

This reduction of the noise in each of the two output channels has long been known by radio astronomers (References 40, 41). It is a mystery to the writer why this case continues to be ignored in the general literature.

The value K = 1, just given for the fully-coherent radiometer case, assumes that there is one input signal that is divided in a hybrid in order to drive the two signal channels. However, often we can use two antennas. For instance, in the case of aperture synthesis there are two antennas at the ends of an interferometer baseline. In general, when probe antennas are used to measure the field in a Fourier plane, these antennas receive power from the entire field of view. In such a case two antennas mounted very close together can both receive power from the entire field-of-view. The two antennas are similar to a monopulse pair such as is used in amplitude monopulse. Use of two antennas doubles signal power. Therefore, K becomes $1/\sqrt{2}$. The treatment of the temperature sensitivity calculation is entirely different, if there is not amplification ahead of the square-law detector. This is the usual state of affairs in infra-red radiometers. This arrangement is also appealing in the microwave region in application where a very simple radiometer is desired, even at the expense of loss of performance. It also can become attractive when a great number of channels must be monitored simultaneously.

In the usual microwave case, where there is amplification ahead of the square-law-detector/multiplier, the performance of a spectrometer using filters, and a Fourier-transform spectrometer, using autocorrelation, is the same. However, when there is no RF amplification, the transform spectrometer temperature resolution is

$$\Delta T_{TRANSFORM} = \sqrt{\frac{2}{N}} \, \Delta T_{FILTER} \qquad (I\text{-}3\text{-}9)$$

where N is the number of frequency bins. The factor $\sqrt{2/N}$ is called Fellgett's Advantage.

The sensitivity of transform radiometers without RF amplification is treated in Appendix A.

We will now summarize the values of K for various types of correlating radiometers in Table 3-3. For comparison purposes the values of K for the common types of power measuring radiometers will be given in a parallel column. Correlating radiometers with separate antenna inputs to the two signal channels will be identified by an asterisk. Transform spectrometers can reduce K by a factor of $1/\sqrt{2}$ by measuring $\phi(\tau)$ over a range of delays which extends from $-\tau_1$ to $+\tau_1$. The values of $\phi(\tau_1)$ are added in the in-phase channel outputs, whilst the value of $\phi(-\tau_1)$ is multiplied by minus one and added to the value of $\phi(\tau_1)$ in the quadrature channel outputs. In both cases the signal add amplitude-wise and noise in the two outputs adds power-wise. The result is the $1/\sqrt{2}$ improvement in K already referred to.

In Table 3-3, FTS will stand for Fourier-transform spectrometer.

TABLE 3-3. VALUES OF THE RADIOMETER FIGURE-OF-MERIT, K

Correlating Types	K	Power-Measuring Types	K
Phase-Switched, Square-Wave Modulation and Demodulation	2	Dicke, Squarewave Modulation and Demodulation	2
Phase-Switched, Square-Wave Modulation and Demodulation*	$\sqrt{2}$	Dicke, Squarewave Mod Sine Wave Demod	$\pi/\sqrt{2}$
Homodyne (Faris' case), Total Power	$\sqrt{2}$	Dicke, Sine Wave Mod and Demod	$2\sqrt{2}$
Thermal-Gradient	$\sqrt{2}$	Dicke, Push-Pull (Graham's) Squarewave Modulation and Demodulation	$\sqrt{2}$
FTS (Positive Delay Only)	$\sqrt{2}$		
FTS (Positive Delay Only)*	1	Thermal Gradient	1
FTS (Symmetric Delay)	1	Total Power	1
Fully-Coherent	1		
Fully-Coherent*	$1/\sqrt{2}$		
FTS (Symmetric Delay)*	$1/\sqrt{2}$		

*Star indicates each correlator channel is fed by a separate antenna.

3.2 RECEIVER SENSITIVITY

In order to calculate the performance of point designs some data on contemporary receiver noise temperatures are needed. Two graphs are given. One by Weinreb (Reference 42) emphasizes FET amplifiers (Figure 3-1). The other, due to Cardiasmenos (Reference 43), shows the performance of super-heterodyne receivers (Figure 3-2). This second paragraph has been redrawn, since the original did not reproduce well.

3.3 GAIN FLUCTUATION NOISE

Power output variations due to receiver gain fluctuation are fully as important as receiver noise in determining the system temperature resolution. To include fluctuation noise in the temperature resolution calculation, observe that for a total power radiometer

$$P = T_{SYS}G \tag{I-3-10}$$

where P is the RF power input to the correlator or square-law detector, as the case may be. The total differential of (I-3-10) is

$$\Delta P = G\delta T_{SYS} + T_{SYS}\,\delta G \tag{I-3-11}$$

G may be considered to be a system calibration factor. So if we divide (I-3-11) by G, we get the correct answer $T_{SYS} = \Delta T$ when G is zero. However, if G is not zero, division by G gives

$$\frac{\Delta P}{G} = \Delta T + T_{SYS}\left(\frac{\delta G}{G}\right) \tag{I-3-12}$$

The second term is error due to gain fluctuation. Since receiver noise and gain fluctuation are both noise-like quantities, their variances add to produce the total error. So

$$\Delta T_{\Sigma} = \left[\Delta T^2 + T_{SYS}^2\left(\frac{\delta G}{G}\right)^2\right]^{1/2} \tag{I-3-13}$$

NOTES:

NOISEFIGURE 10 LOG F, AND NOISE TEMPERATURE $T_n = 290°$ (F-1), VERSUS FREQUENCY FOR VARIOUS 1980 STATE-OF-THE-ART LOW-NOISE DEVICES. THE 300 K BIPOLAR TRANSISTOR, FET, AND PARAMP VALUES ARE TAKEN FROM MANUFACTURERS DATA SHEETS (1) — (3), THE 20 K FET CURVE IS FROM THE DATA OF THIS PAPER PLUS DATA OF OTHERS (4) — (6) AT 0.5, 1.4, AND 12 GHz, RESPECTIVELY. THE 20 K PARAMP, 4 K MASER (INCLUDING PARAMETRIC UP-CONVERTER AT LOWER FREQUENCIES), AND 20 K MIXER RESULTS (WHICH ARE SSB AND INCLUDE IF NOISE) ARE FROM SYSTEMS IN USE AT NATIONAL RADIO ASTRONOMY OBSERVATORY (NRAO). THE NATURAL NOISE LIMITATIONS DUE TO GALACTIC NOISE, THE COSMIC BACKGROUND RADIATION, AND ATMOSPHERIC NOISE ARE FOR OPTIMUM CONDITIONS AND ARE TAKEN FROM (9) PLUS POINTS AT 22 GHz AND 100 GHz MEASURED AT NRAO.

FIGURE 3-1. PERFORMANCE OF LOW NOISE DEVICES
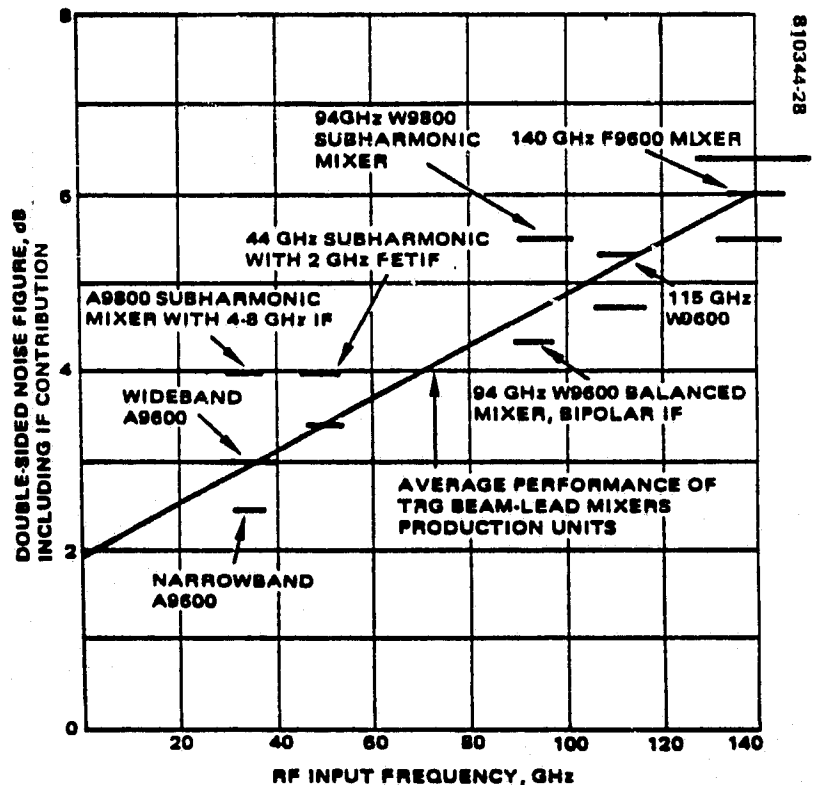
FIGURE 3-2. PERFORMANCE OF TRG SUPERHETS

In a Dicke radiometer the unknown and reference temperatures are both subject to the same gain fluctuation. Therefore the gain fluctuation error is $|T_{SIG} - T_{REF}|$. If the difference between T_{SYS} and T_{REF} is servoed to zero, the gain fluctuation error also goes to zero.

In the correlating radiometer case receiver noise in the two receiver channels does not correlate when the correlation function is generated. The autocorrelation functions of the receiver noises appear at DC and are rejected by the bandpass filter between the square law detector and the I&Q phase detectors.

Of course, if the radiometer is operated in a total-power mode, the noise from both channels will not be rejected and will appear in the radiometer output as a pedestal upon which the signal rides. Then gain fluctuation will produce just as large errors as in the noncoherent, total-power radiometer case.

In the frequency offset mode the correlating radiometer will still experience gain fluctuation errors due to variable amplification of the signal itself. However the signal is usually smaller than the receiver noise (e.g., 200°K vs. 1000°K). So the gain fluctuation noise is reduced accordingly.

However, it will be found when calibration is discussed, that a reference signal can be injected into the two correlator channels in a manner that causes the correlator output to be proportional to $(T_{SIG} - T_{REF})$. Therefore if this difference is driven to zero, gain fluctuation noise will go to zero. This behavior is very similar to the case of a Dicke with a servoed reference temperature.

4.   CALIBRATION

4.1   CLAPP CALIBRATOR

Almost all of the calibration methods used by noncoherent radiators can be adapted for use in a correlating radiometer. However, a calibration method proposed by Clapp (Reference 44) is clearly superior to other ways of calibrating coherent radiometers. This method can only be used this radiometer type. A functional block diagram of this calibration method, taken from Clapp and Maxwell's paper is shown in Figure 4-1.

Following Clapp, the inputs to amplifiers G_1 and G_2 are

$$\begin{cases} S_1 = A_1 + (B_{pr} + B_{nr} + B_{pi} + B_{ni}) \\ \\ S_2 = A_2 - (B_{pr} - B_{nr} - jB_{pi} + jB_{ni}) \end{cases} \qquad (I-4-1)$$
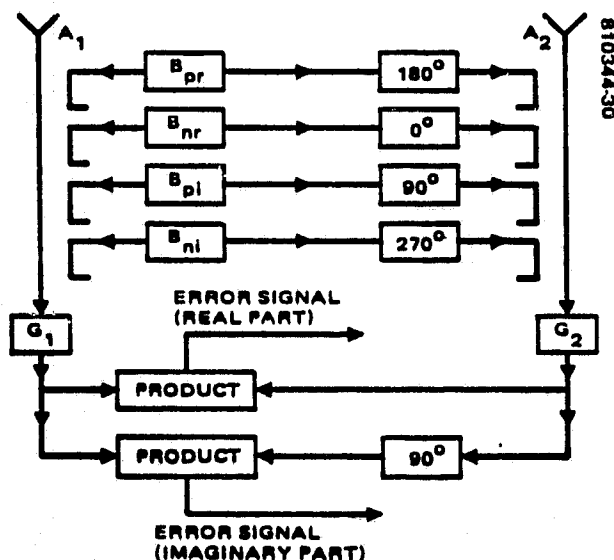
FIGURE 4-1. SYSTEM CONCEPT, COMPLEX CORRELATION RADIOMETER

The antenna signals are

$$\begin{cases} A_1 = A_{1r} + jA_{1i} \\ \\ A_2 = A_{2r} + jA_{2i} \end{cases} \tag{I-4-2}$$

The cross-correlation function F can be defined by

$$F = A_1 A_2^* = F_r + jF_i \tag{I-4-3}$$

Its real and imaginary parts are given by

$$\begin{cases} F_r = (A_{1r}A_{2r} + A_{1i}A_{2i}) \\ \\ F_i = (A_{1i}A_{2r} - A_{1r}A_{2i}) \end{cases} \tag{I-4-4}$$

A complex error signal will be developed by the pair of product detectors shown in Figure 4-1. This error signal has the form

$$E = (G_1 S_1)(G_2 S_2)^* = (G_1 G_2^*)(F_r - H_r) + j(G_1 G_2^*)(F_i - H_i) \tag{I-4-5}$$

H is the contribution of the complex noise source whose real and imaginary parts are

$$\begin{cases} H_r = B_{pr}B_{pr}^* - B_{nr}B_{nr}^* \\ \\ H_i = B_{pi}B_{pi}^* - B_{ni}B_{ni}^* \end{cases} \tag{I-4-6}$$

Cross product terms which average to zero have been dropped.

The real part of the error signal actuates a servo loop which controls noise sources B_{pr} and B_{nr}. Likewise the imaginary part of E controls B_{pi} and B_{ni}. When these servo loops have been brought to equilibrium the noise sources will be held at levels which can be seen to be

$$\begin{cases} F_r = H_r = |B_{pr}|^2 - |B_{nr}|^2 \\ \\ F_i = H_i = |B_{pi}|^2 - |B_{ni}|^2 \end{cases} \qquad (I-4-7)$$

In a correlator which measures only a single temperature only the noise sources B_{pr} and B_{nr} are needed.

4.2 HACH SERVO LOOP

Hubbard and Erickson (Reference 45) used thermionic noise diodes as noise sources when they mechanized the radiometer analyzed by Clapp and Maxwell. However, the Hach servo loop discussed in Section 2, which uses semiconductor noise diodes, is better for orbital use. The phase tracking between the two channels may be quite poor without preventing the servo loops from reaching equilibrium. Let ϕ_1 and ϕ_2 be the cumulative phase shifts through the two channels from the input antennas or input power-dividing hybrid. Then $(\phi_1 - \phi_2)$ is the relative phase shift between channels. It is a measure of phase tracking.

Clapp and Maxwell computed the convergence of the servo loops and plotted their results on the complex plane. Their plots, showing servo convergence for relative phase shifts of +45° and −84°, are reproduced here as Figure 4-2. In actuality the relative phase shift will not exceed 10° to 20°, for a well designed radiometer, under the worst of conditions.

Figure 4-3 shows experimental results of Hubbard and Erickson's (Reference 45) loop stability tests. Part (a) shows the radiometer response with various relative phase shifts. At +84° one can see an increase in servo noise. At +90° the servo breaks into oscillation as one would expect. Part (b) shows the result of varying servo gain ±5 db around a reference gain of 0°. Again the system input is 1200°K. Only slight variations in system response can be seen. Figure 4-4 shows receiver response vs correlated signal input.
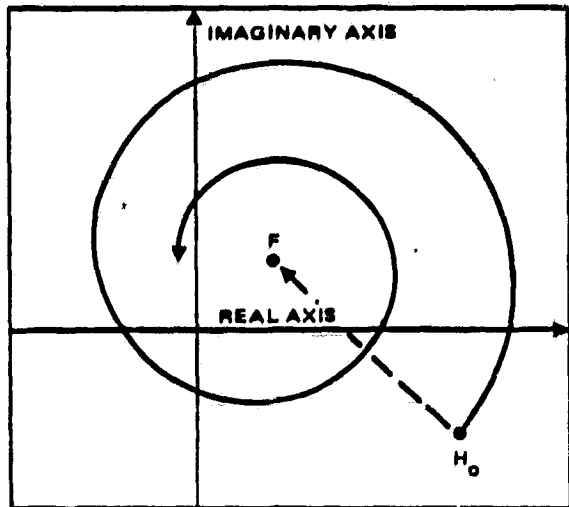
NOTE: CONVERGENCE OF COMPLEX NOISE SOURCE,
H, FROM INITIAL VALUE, H_0, TOWARD FINAL VALUE, F

a) THE SOLID CURVE IS THE SPIRAL FOLLOWED WHEN THE RELATIVE PHASE SHIFT OF THE TWO AMPLIFICATION CHANNELS IS +45°. THE DASHED LINE SHOWS THE CONVERGENCE WHEN THE RELATIVE PHASE SHIFT IS 0°
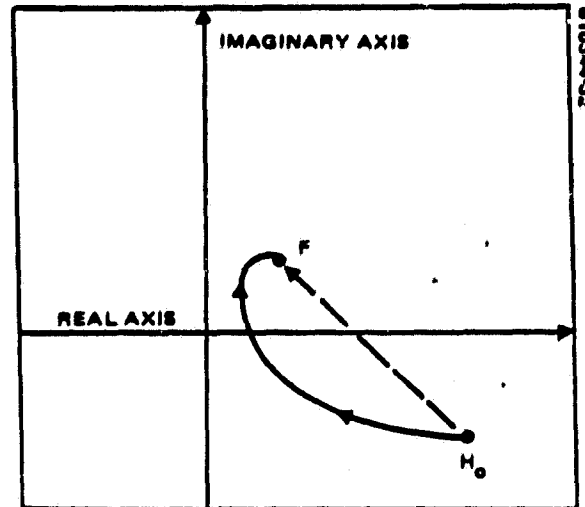
b) THE SOLID CURVE IS THE SPIRAL FOLLOWED WHEN THE RELATIVE PHASE SHIFT OF THE TWO AMPLIFICATION CHANNELS IS -84°. THE DASHED LINE SHOWS THE CONVERGENCE WHEN THE RELATIVE PHASE SHIFT IS 0°

FIGURE 4-2. SERVO CONVERGENCE
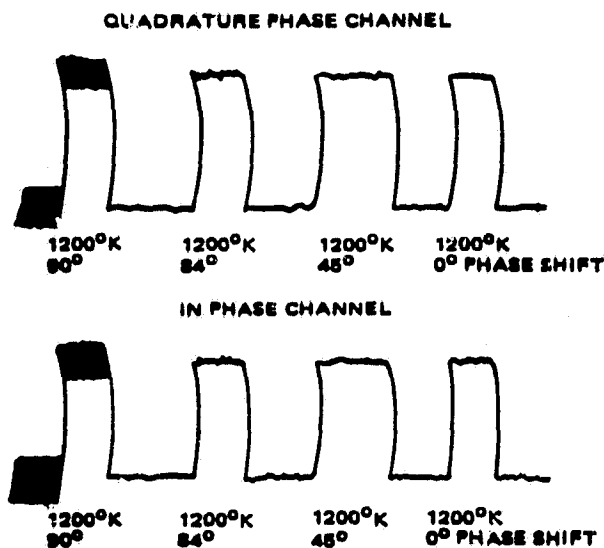
a) RECEIVER RESPONSE VERSUS RELATIVE AMPLIFIER PHASE SHIFT
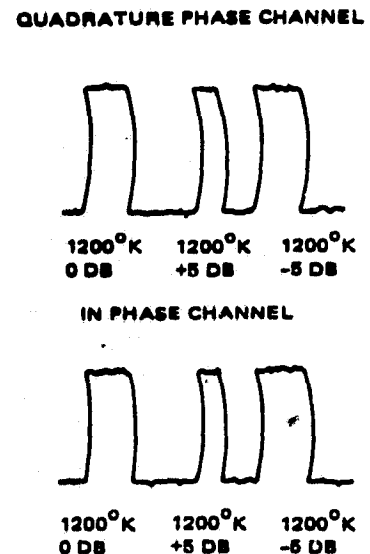
b) RECEIVER RESPONSE VERSUS AMPLIFIER GAIN CHANGE

FIGURE 4-3. RECEIVER RESPONSE
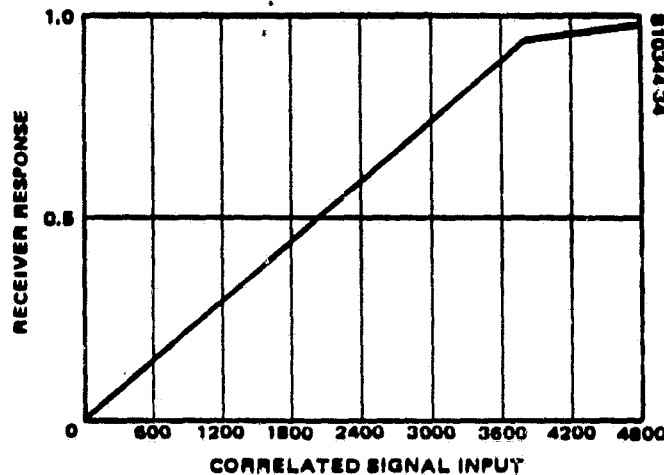
FIGURE 4-4. RECEIVER RESPONSE VERSUS
CORRELATED SIGNAL INPUT

4.3 REFERENCE NOISE-SOURCE STANDARDS

In the past noise diodes and gas-discharge tubes have been popular
standard noise sources (Figure 4-5). The thermionic noise diode will not be
considered further. A complete treatment of temperature-limited vacuum tube
slot noise is in the Rad Lab Series (Reference 46).

Gas-discharge sources have been widely used. They operate at frequencies
up to tens of gigahertz. The effective temperature is very close to the kinetic
temperature of the plasma electrons in the discharge. This is of the order
of 10,000°K. The most common gas used is argon, although neon has been used
at times. Again these tubes belong mainly in the past.
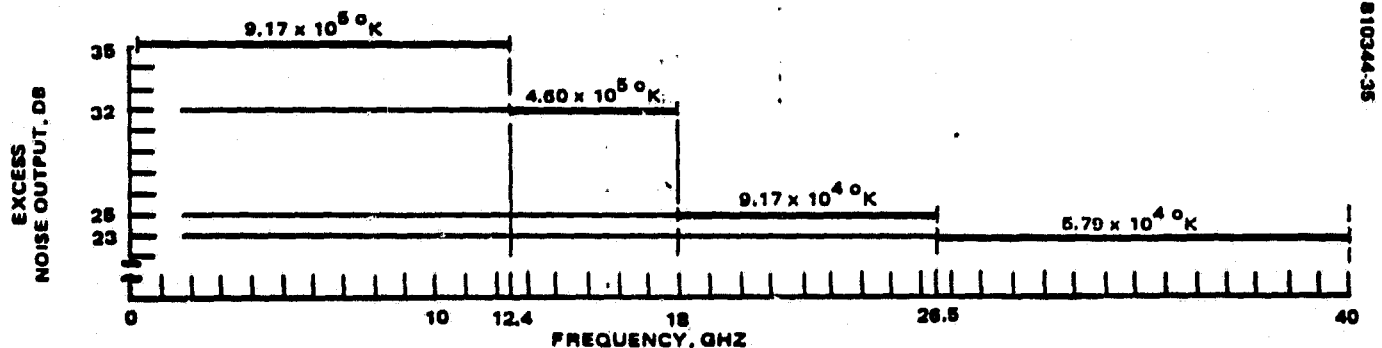
FIGURE 4-5. SYSTEM NOISE SOURCES – TYPICAL NOISE OUTPUT

At the present time the most popular standard noise source is the semiconductor noise diode. Microwave Semiconductor Corp. is a supplier of these devices. ·Some pertinent data follows abstracted from their literature.

General operating characteristics:

Operating temperature	-55° to 100°C
Temperature sensitivity	0.01 db/°C
Voltage sensitivity	0.1 db/%
Current sensitivity	0.03 db/%
Switching speed	1.0 μsecond

The highest frequency diode is quite flat in response. It appears that it can be used out to 100 GHz or so.

With proper temperature control and good stabilization of applied power the semiconductor noise diode can hold calibration for long periods of time. However, calibration against an absolute standard is easier to provide than for a noncoherent radiometer. The reason lies in the fact that a correlating radiometer, using a hybrid power splitter at its input, measures the difference between the sum and difference ports of the hybrid shown in Figure 4-6.

When calibration is required the switch-pair are thrown to "calibrate" position. The servo loop then produces a frequency which sets the diode noise output waveform to a value such that the 200°C temperature difference between the 100°C and 300°C loads is balanced out by the diode noise injected into the mixers in the calibration system by directional couplers 3 and 4.

When the switch-pair returns to "measure" position, the servo loop balances out the unknown temperature difference between the input and the 50°C load on the hybrid difference channel. It may be seen that the ratio of the unknown difference, present in the measure mode, and the known difference, present in the calibrate mode, is the same as the ratio between noise waveform duty factors present in two modes. So the unknown temperature difference is this ratio times the temperature difference between the calibration loads. Adding the temperature of input hybrid delta to the calculated difference gives the unknown temperature.
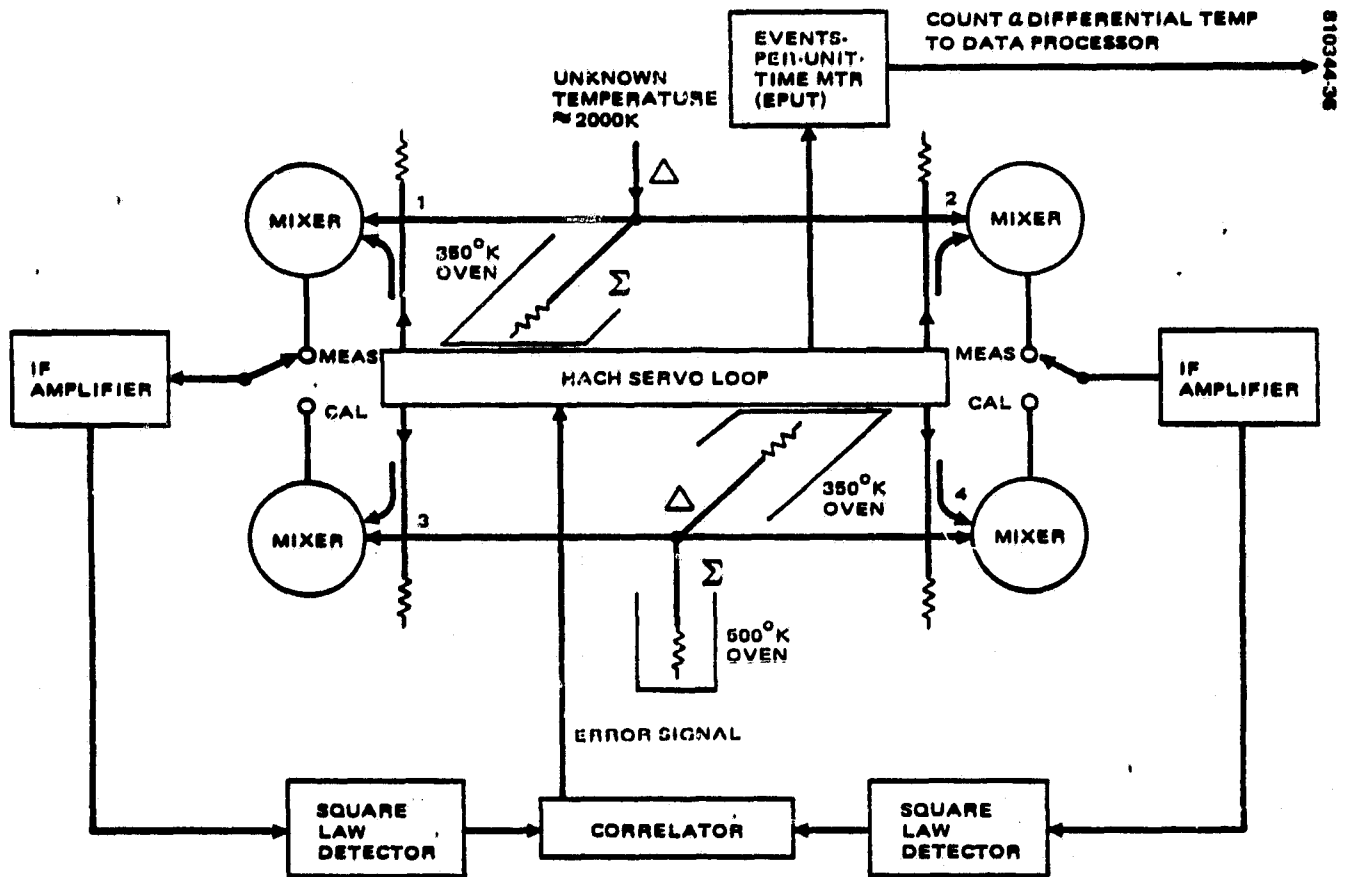
FIGURE 4-6. CALIBRATION OF NOISE DIODE AGAINST AN ABSOLUTE STANDARD

4.4   DISCUSSION OF THE CALIBRATION CHARACTERISTICS OF FREQ-DOMAIN IMAGERS

If the radiometer measures the autocorrelation of the signal, the use of a servo loop of type just described remedies a fault of autocorrelators, the large dynamic range they require. This results because the zero differential delay bin of the autocorrelator measures the total signal power. When the signal has a wide bandwidth, as most radiometric signals do, the autocorrelation function may drop 10 or 20 db in the first couple of delay bins adjacent to the $\tau = 0$ bin. As a result a large portion of the system dynamic range is used up just to accommodate the $\tau = 0$ bin.

The servo reduces the output of the $\tau = 0$ bin to zero. This greatly reduces the required dynamic range. The source spectrum autocorrelation function is subtracted from the signal autocorrelation function not only in the $\tau = 0$ bin but also in all the other bins holding the autocorrelator output.

If the noise-source spectrum is flat, the autocorrelator output will have the form shown in (b) of Figure 4-7. To left of (b) diagram the original autocorrelation is shown before subtraction of the noise-source autocorrelation. The spectrum of the unknown signal is shown above the two correlation functions.

·After (b) is Fourier-transformed the known noise-source spectrum is added in to recover the signal spectrum.

Very often the average shape of the unknown spectrum is known. The fluctuations of actual spectrum around the average may be smal compared to the average spectrum power. This is true in the case of temperature sounding. The spectrum calculated from the standard atmospheric temperature profile is the average spectrum. Its power density expressed as a temperature, is in the range of several hundred degrees Kelvin. The rms fluctuation of actual spectrums about the average spectrum never exceeds 5°K. Assume that the system can accommodate 3σ variations in the spectrum. We find that the ratio of average spectrum to spectrum fluctuations is about 210°K/30° = 7°, 17 db.

In cases of this sort we can achieve considerable improvement in system performance by shaping the noise source spectrum before its output is injected into the radiometer. If it is shaped to resemble the average spectrum and subtracted from the signal spectrum, the residual will be much smaller than the total spectrum. This is shown in Figure 4-8 where the total and residual spectrums are sketched together with their correlation spectrums.
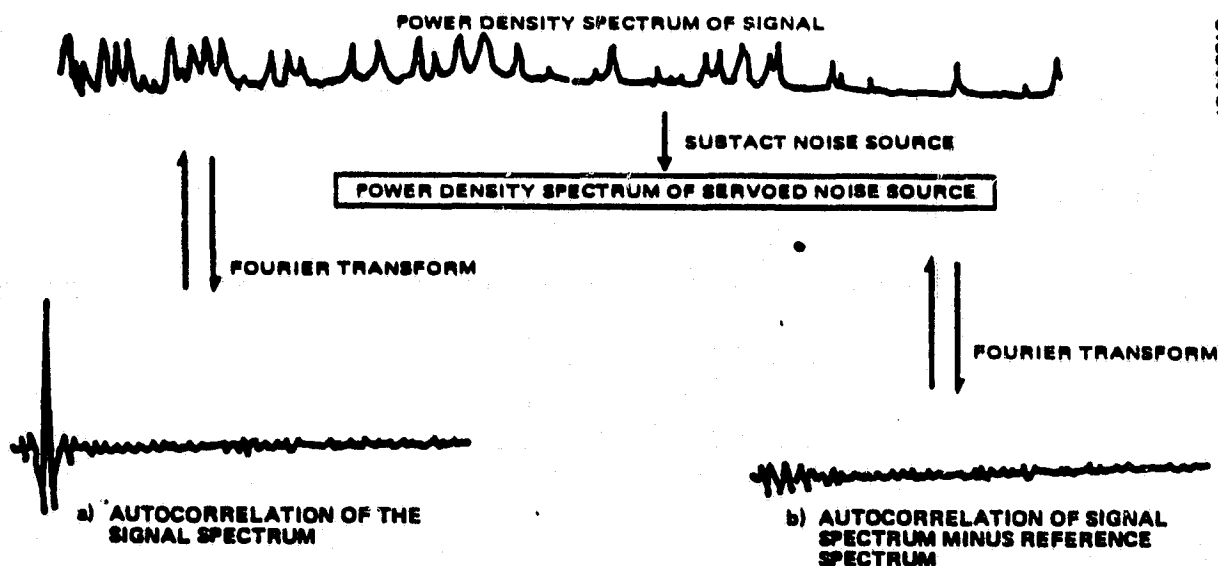
FIGURE 4-7. MODIFICATION OF OUTPUT AUTOCORRELATION BY REFERENCE NOISE SPECTRUM
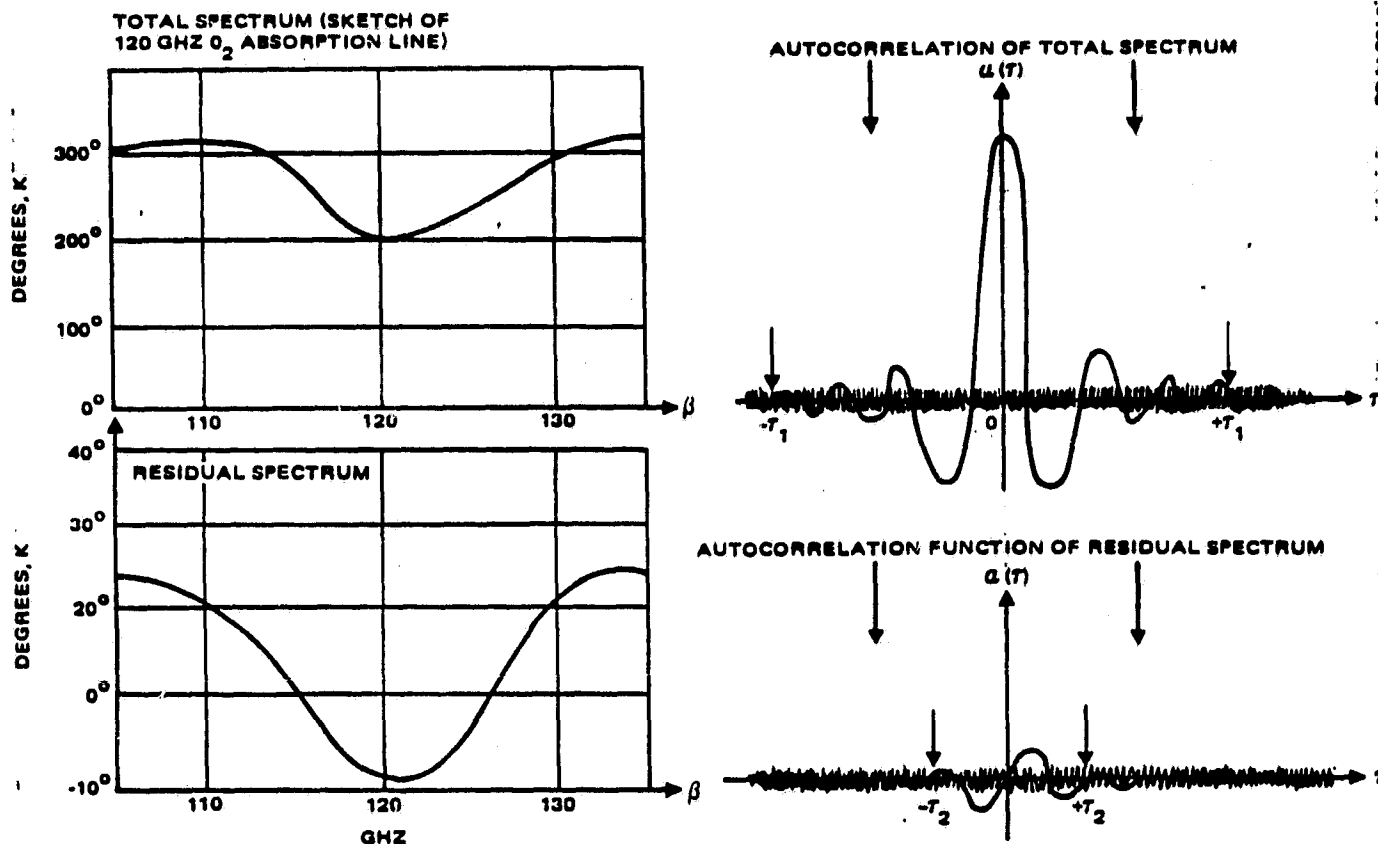
FIGURE 4-8. TOTAL AND RESIDUAL SPECTRUMS AND THEIR AUTOCORRELATION FUNCTIONS

It may be seen that since the residual spectrum has much less power in it than the total spectrum, its correlation function will reach the noise level of the radiometer for a smaller differential delay than the value where $\phi(\tau)$ of the total spectrum reaches zero. It is obvious that correlation measurements should be truncated for a value of τ where $\phi(\tau)$ subsides into the noise. Beyond that point more noise would be introduced into the transform computation than signal data.

Elimination of the delay bins in the τ_1 to τ_2 regions makes the auto-correlator smaller and simpler.

Elimination of the bins in the τ_1 to τ_2 region also reduces the noise introduced into the output spectrum. This reduction can be an order of magnitude or so, since the reduction in spectrum power is an order of magnitude or more. The net result is that the noise variance of the system tends to be proportional to the real unknown spectrum, namely the residual spectrum. Nothing is lost. We can get the total spectrum by adding the average spectrum back into the residual spectrum. We have used our a prior knowledge about the spectrum to reduce noise variance.

Even in mapping, where absolute temperature values are not important, removal of the average temperature can improve performance. A block diagram of a servo loop to accomplish this is shown in Figure 4-9.

The loop is biased so that increase in error signal reduces the injected reference. The composite signal formed by injecting the noise source into the signal channel with negative polarity will adjust itself so that a few negative going excursions will hold the loop in equilibrium as shown in Figure 4-10.

This arrangement permits display of the signal with maximum dynamic range. Dynamic range is always in short supply in display systems. Signal control which produces better image dynamic range always causes dramatic improvements in image quality.

When precision measurements are required the elimination of the sky-horn or cryogenic load as the low temperature reference is important. Elimination of the sky-horn makes it possible to design better radiometer installations aboard spacecraft. When power-measuring radiometers are used, the sky-horn
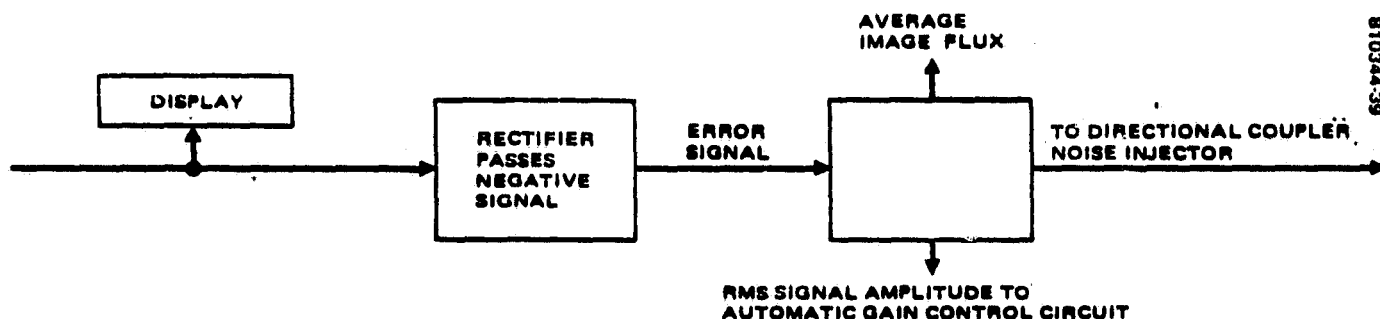
FIGURE 4-9. REFERENCE NOISE INJECTION SERVO LOOP FOR TERRAIN MAPPING
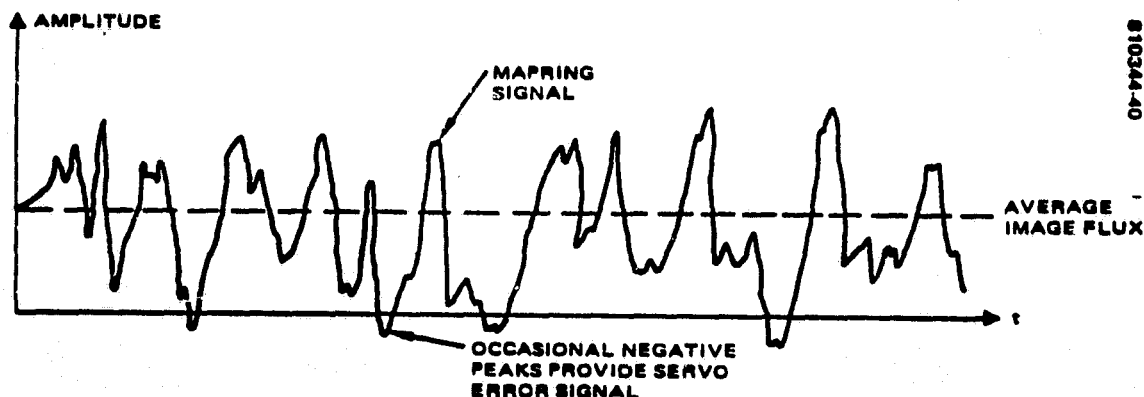
FIGURE 4-10. STABILIZED MAPPING VIDEO

must see the sky well above the horizon in order to produce a predictable temperature.   This often drives the designer to put the radiometer antenna on top of the spacecraft. Now the scanning beam must look down outboard from the craft — peering over the edge as it were.

This causes restrictions on field-of-view, since the beam cannot look through the spacecraft.

What is worse, the top mount causes undesired interactions between antenna and spacecraft structure. If one visualizes the beam as simply emerging from the antenna and continuing to the ground, it is hard to see why the interaction occurs.

However, the beam forms out of a haze of partially reactive electromagnetic field extending out to about d^2/λ. It is this haze which couples to the antenna surroundings. Rapid changes in the antenna pattern can make it very hard to correct for antenna pattern bias by deconvolution.

It is also possible to get beam dropouts in directions where there is tight antenna/structure coupling. As a result, most aircraft antenna installations are checked out on model ranges in order to eliminate unpleasant surprises.

The Clapp calibrator makes the sky-horn unneeded. Referring to Figure 4-1 we see that all that is required to get a zero or negative temperature is the reduction of B_{pr} and increase in B_{nr} until sufficient negative temperature has been injected into the system to produce the desired condition. Since there is no sky-horn, the scanning antenna can be mounted on the bottom of the spacecraft where field-of-view problems are reduced. Coupling of the antenna field to the spacecraft is also much reduced.

There is one other characteristic of the correlating radiometer which is important enough to warrant discussion. In correlating radiometer designs where correlation functions are formed, this function will be Fourier transformed to obtain a spectrum. If the function is a cross correlation formed by aperture synthesis, the transform will be an image, not a spectrum.

In either case, the transform calculation produces noise in the spectrum bins or in the image pixels which originates in all the delay bins of the correlator. The transform process consists of multiplications of the values by phase shifting exponentials of the form, $\exp(-j2\pi\phi)$. The phase shifted terms, when summed, form the sequence of value output by the transform algorithm. The noise in the input bins comes along and also is multiplied by the exponentials. However, being noise, this has no effect. The summing steps in the transform then adds all the noise samples power-wise.

As a result the transform algorithms acts like a food blender. All output resolvable elements have the same noise mix from the correlator delay bins. Each signal element output by the transform algorithm is a mixture of all $\phi(\tau)$ values. This process ties the whole instrument together giving it one response. Variations in response or noise level in the autocorrelator bins do not appear in the spectrum or pixel bins. One calibration operation calibrates the total output. This ensures smooth spectrums and images devoid of streaking and uneven response.

This property of the transform is so valuable it is purposely built into some image transmission systems. The image is transformed to the frequency domain before transmission and converted back on reception. If interference or channel drop-outs produce glitches in the spectrum during transmission, the result after inverse transformation on reception is a slight drop in the image signal-to-noise ratio distributed throughout the image. This is much more desirable than the defects in the image itself, if it were transmitted without transformation.

5. SPATIAL FREQUENCY CONCEPT

5.1 INTRODUCTION

Almost from the beginning of electronic technology amplifiers, transmission lines and other devices have been analyzed by calculating their behavior in the frequency domain. This involves finding the spectrum of the applied signal. Modification of the spectrum by passage through the device is then calculated. The properties of the device regarded as a filter are the characteristic of interest in doing this analysis.

When thinking about antennas, the concept of the antenna as a filter is equally useful. In the case of the electrical filter, mentioned first, the signal is a function of time. It is expanded into a spectrum of temporal frequencies. In the case of antennas and lenses the signal is a two-dimensional function whose independent variables are two angles. This function can be transformed into a two-dimensional spectrum whose independent variables are two spatial frequencies.

5.2 HEURISTIC DISCUSSION OF SPATIAL FREQUENCY

The spatial frequency spectrum describes the distribution of some property over the field-of-view of the antenna/lens. For instance, let the property be radiant brightness. Figure 5-1(a) shows a one-dimensional version of brightness distribution. Let us treat this distribution just as if it were a function of time, not angle. Let its Fourier transform be found. The new independent variable is a spatial frequency in cycles per radian or cycles per degree and the new function is a spatial frequency spectrum. Figure 5-1(b) shows a sinusoidal component of the brightness distribution present in 5-1(a). It corresponds to a small portion of the total spatial frequency spectrum which falls within a small frequency increment. At zero angle it may be seen that the sinusoidal wave representing this component is neither at zero nor at a maximum. This means the waveform is neither a pure sine wave nor a pure cosine wave. However, it can be represented as the sum of a sine wave and a cosine wave, each of proper amplitude. The cosine component is a part of the real spectrum. The sine component is part of the imaginary spectrum. If the spectrum is observed with a fully coherent radiometer, the real spectrum is the portion of signal which passes through the in-phase synchronous demodulator (phase-detector). The imaginary component is recovered from the total signal by the phase detector in quadrature with the reference.
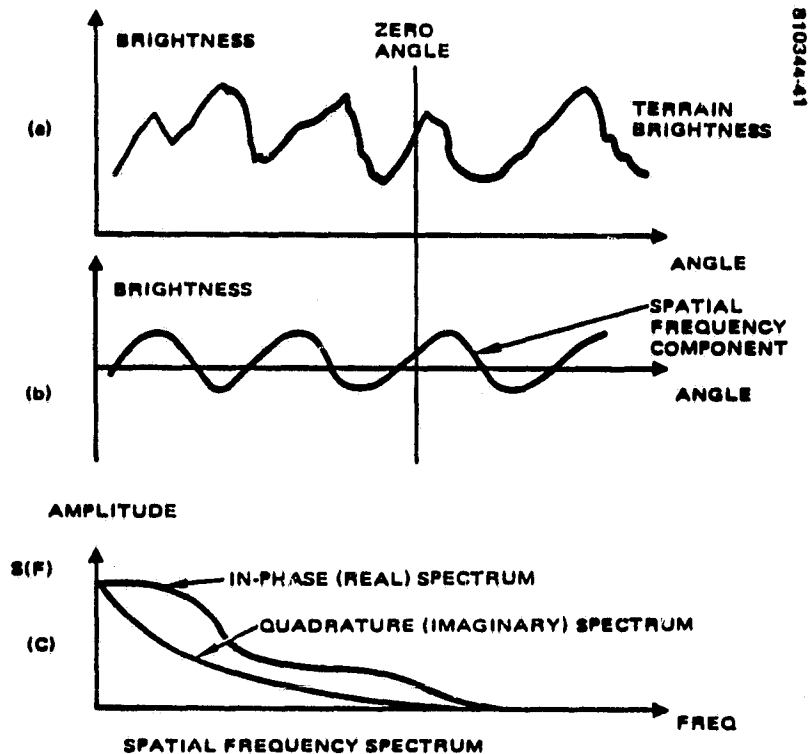
FIGURE 5-1. RELATION OF THE SPATIAL FREQUENCY
SPECTRUM TO SCENE BRIGHTNESS DISTRIBUTION

Note that (c) is an amplitude distribution. The radiant brightness in $^{\circ}K$ is proportional to rf power density (watts/Hz/steradian). However, a correlator, like a square law detector, outputs a voltage proportional to rf power level. So the video amplitude is proportional radiant brightness and the video variance has the dimensions of degrees Kelvin squared ($^{\circ}K^2$). This point is important because the amplitude spectrum is just another form of the image signal. An inverse transform will recover the brightness distribution. The power density spectrun, $S(F)S*(F)$, has lost the phase information. So the scene brightness distribution cannot be recovered. In short, the power density spectrum is a piece of statistical information about many scenes, while the amplitude spectrum of a given scene is that scene in a coded form resulting from Fourier transformation.

5.3   DETERMINATION OF THE POWER-DENSITY SPECTRUM OF THE SCENE

The statistical data stored in the video power density spectrum is of first importance in designing radiometers and predicting their performance. This will become apparent when we consider frequency-domain imagers in detail.

At present, data about radiometric-scene statistics is very sparse. In order to have a reasonable spectrum for use in this study, estimates of the video power density spectrum were made by three different methods. The details

of the process of making these estimates are given in Appendix C, "Microwave Radiometer Image Statistics."

The first spectrum estimate is based on the work of Mandelbrot (Reference 48). His spectrums have the form $P(F) = Const. F^{-2H-1}$. These spectrums are associated with scenes whose statistics are independent of scale. This is a property of many natural scenes. Let these scenes be observed over a range of magnifications. The resulting scenes are said to be self-similar with respect to a set of ratios r_i when the statistics of all the scenes are the same.

A spectacular example of this behavior was the series of images of Jupiter sent back by Voyager as it closed to a shorter and shorter range. As range decreased Jovian weather patterns will become too big to be seen as separate entities in the image. However, little pattern details which were not resolved before now could easily be seen. The new patterns had the same overall appearance as the old patterns.

Mandelbrot's final choice for H was arrived at by comparing the properties of statistically-generated landscapes with actual landscapes. These computer generated landscapes are shown on pp 210-215 of reference 48. His final choice for the value of H is 3/4.

Another means of determining H is based on the fact that the spectrum shape is a measure of redundancy in the scene. If the spatial frequency spectrum were flat, each new pixel value would be completely unpredictable from knowledge of past pixel values. Let the high frequencies in the spectrum be reduced. As this reduction increases and the spectrum drops off more and more steeply redundancy increases.

There are ways of coding the image which decrease the bandwidth required to transmit images at a given rate by removing the image redundancy. The amount of image compression obtainable depends on the redundancy of the original image. This in turn depends on the value of H in the general spectrum given for $P(F)$.

In Appendix C.2 it was estimated that satisfactory image compression codes can reduce images with a dynamic range of 1,000 to formats with an information density of one bit per pixel. The value of H which agrees with these facts was then found to be 0.70.

A third approach used experimental microwave radiometer spectrum data supplied by Mr. Charles Hawthorne of the Navy Ordnance Test Station, Inyokern. A regression algorithm was used to fit the experimental data to the power-law spectrum form. The value of H obtained in this manner was 0.60. Figure 5-2 shows the raw spectrum data.

Extension to the actual two-dimensional scene spectrum is made by expressing the spectrum in polar coordinates. The one-dimensional spectrum is then used as the spectrum function along coordinate radials. No variation
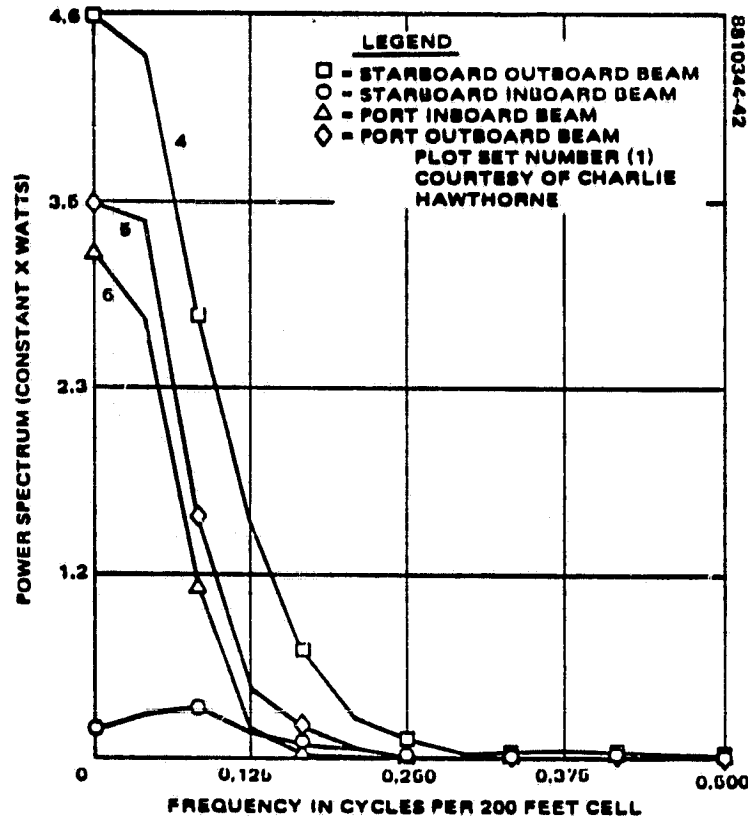
FIGURE 5-2. POWER SPECTRUM VERSUS FREQUENCY

in spectrum value occurs in angle. A value of H = 3/4 was adopted for this study.

5.4 ALTERNATE STATISTICAL DESCRIPTIONS

Other power-density spectrums, as well as alternate statistical formats, have been used to describe scene statistics. They are reviewed in this section.

Viterbi and Omara (Reference 50) treat the case where the scene is isotropic and has a space correlation function which depends only on the total distance between two points in the two-dimensional space. That is

$$\tilde{\phi}(r_x, r_y) = \tilde{\phi}\left(\sqrt{r_x^2 + r_y^2}\right) = \tilde{\phi}(r) \tag{I-5-1}$$

Changing to polar coordinates, let

$$r_x = r\cos\theta_r, \quad r_y = r\sin\theta_r$$

and

$$\omega_x = \omega\cos\theta_\omega, \quad \omega_y = \omega\sin\theta_\omega$$

Then the polar spectrum is

$$\phi(\omega,\theta_\omega) = \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} \phi(r_x,r_y)e^{-i\omega(r_x\cos\theta + r_y\sin\theta)}dr_x dr_y$$

$$= \int_{-\infty}^{+\infty}\int_{0}^{2\pi} \overset{\sim}{\phi}(r)e^{-i\omega r(\cos\theta_r\cos\theta_\omega + \sin\theta_r\sin\theta_\omega)}r\,dr\,d\theta$$

Since there is no θ_ω dependence,

$$\phi(\omega) = \int_{0}^{\infty} \overset{\sim}{\phi}(r)\left[\int_{0}^{2\pi} e^{-i\omega r\cos(\theta_\omega-\theta_r)}\right]r\,dr$$

$$(I-5-2)$$

$$= 2\pi\int_{0}^{\infty} r\overset{\sim}{\phi}(r)J_0(\omega r)\,dr$$

Viterbi and Omara also report a spectrum used by Sakrison and Algazi for television image analysis. It is

$$\psi(\omega) = \frac{2\omega_0}{\omega_0^2 + \omega^2} \qquad\qquad (I-5-3)$$

The corresponding autocorrelation is

$$\phi(r) = \exp(-|r|d) \qquad\qquad (I-5-4)$$

where d is the coherence length of the field.

Berger (Reference 51) uses the spectrum

$$\phi(\omega) = A_n\left[1 + \left(\frac{\omega}{\omega_0}\right)^{2n}\right], \quad n = 1, 2, \ldots\ldots \qquad\qquad (I-5-5)$$

where A_n is chosen so that $(2\pi)^{-1}\int \phi(\omega)d\omega = \sigma^2$.

Pratt (Reference 52) assumes a Markovian source (memoryless source) whose elements have equal variance, σ^2 and which possess adjacent pixel correlation factors of ρ_R and ρ_C along rows and columns.

5.5 SPATIAL FREQUENCY VERSUS TEMPORAL FREQUENCY FILTERS

It should be obvious at this point that an antenna acts as a filter of spatial frequency spectrums in the same manner as electrical filters operate on temporal frequency spectrums.

A detailed development of this parallel between temporal and spatial filters is given by O'Neill (Reference 53). Their parallel properties are shown in O'Neill's Table 2-1, which is reproduced here as Table 5-1. The quantities in Table 5-1 are defined in Table 5-2.

TABLE 5-1. SPATIAL FILTERS VERSUS TEMPORAL FILTERS

	Time filters	Space filters
Superposition integral	$y(t) = \int_{-\infty}^{t} h(t - t')x(t') \, dt'$	$i(x, y) = \iint s(x - \xi)o(\xi) \, d\xi$
Periodic input	$x(t) = \sum_{-\infty}^{+\infty} X_n e^{i\omega_n t}$	$o(\xi, \eta) = \sum_{m,n=-\infty}^{+\infty} O_{mn} e^{i\rho \cdot \omega_{mn}}$
	$y(t) = \sum_{-\infty}^{+\infty} Y_n e^{i\omega_n t}$	$i(x, y) = \sum_{m,n=-\infty}^{+\infty} I_{mn} e^{ir \cdot \omega_{mn}}$
	$Y_n = H(\omega_n)X_n$	$I_{mn} = \tau(\omega_{m,n})O_{mn}$
Transient input	$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(\omega)e^{i\omega t} \, d\omega$	$o(\rho) = \frac{1}{(2\pi)^2} \iint O(\omega)e^{i\omega \cdot \rho} \, d\omega$
	$y(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} Y(\omega)e^{i\omega t} \, d\omega$	$i(r) = \frac{1}{(2\pi)^2} \iint I(\omega)e^{i\omega \cdot r} \, d\omega$
	$Y(\omega) = H(\omega)X(\omega)$	$I(\omega) = \tau(\omega)O(\omega)$
Random input	$\phi_{xx}(\tau) = \langle x(t)x(t + \tau)\rangle$	$\phi_{oo}(\rho') = \langle o(\rho)o(\rho + \rho')\rangle^*$
	$\phi_{yy}(\tau) = \langle y(t)y(t + \tau)\rangle$	$\phi_{ii}(\rho') = \langle i(\rho)i(\rho + \rho')\rangle$
	$\Phi(\omega) = \int_{-\infty}^{+\infty} \phi(\tau) \cos \omega\tau \, d\tau$	$\Phi(\omega) = \iint \phi(r)e^{-i(\omega \cdot r)} \, dr$
	$\phi_{yy}(\tau) = \int_{-\infty}^{+\infty} \phi_{hh}(t - \tau)\phi_{xx}(t) \, dt$	$\phi_{ii}(r) = \iint \phi_{ss}(\rho - r)\phi_{oo}(\rho) \, d\rho$
	$\Phi_{yy}(\omega) = |H(\omega)|^2 \Phi_{xx}(\omega)$	$\Phi_{ii}(\omega) = |\tau(\omega)|^2 \Phi_{oo}(\omega)$
	$\phi_{xy}(\tau) = \int_{-\infty}^{+\infty} h(\tau + t)\phi_{xx}(t) \, dt$	$\phi_{io}(r) = \iint s(r + \rho)\phi_{oo}(\rho) \, d\rho$
	$\Phi_{xy}(\omega) = \Phi_{xx}(\omega)H(\omega)$	$\Phi_{io}(\omega) = \tau(\omega)\Phi_{oo}(\omega)$

TABLE 5-2. DEFINITIONS OF FILTER FUNCTIONS

Time		Space	
$x(t)$	= input waveform	$O(\xi,n)$	= scene brightness distribution
$y(t)$	= output waveform	$i(x,y)$	= image brightness distribution
$h(t-t')$	= filter memory function	$S(x-\xi,y-n)$	= lens spread function
X_n	= Fourier coefficients of $X(t)$	O_{mn}	= Fourier coefficients of
Y_n	= Fourier coefficients of $Y(t)$	I_{mn}	= Fourier coefficient of
ω_n	= harmonic frequencies	$\vec{\omega}_{mn}$	= radius vector of harmonic frequencies
$H(\omega_n)$ or $H(\omega)$	= filter transfer functions	$\tau(\vec{\omega}_{m,n})$ or $\tau(\vec{\omega})$	= modulation transfer functions
$H(\omega_n), H(\omega) \rightleftharpoons h(t)$		$o(\vec{\omega})$	= object spectrum
$X(\omega)$	= input spectrum	$I(\vec{\omega})$	= image spectrum
$Y(\omega)$	= output spectrum	$O(\vec{\rho})$	= scene brightness at $\vec{\rho}$
$\phi_{xx}(\tau)$	= autocorrelation of $x(t)$	$i(\vec{r})$	= image brightness at $\vec{r}$
ω	= radian frequency	$\phi_{oo}(\vec{\rho'})$	= object autocorrelation
t	= time	$\vec{\omega}$	= radius vector $= k\omega_x + \ell\omega_y$
τ	= differential time delay	$\vec{\rho}$	= radius vector in object plane
$\phi_{yy}(\tau)$	= autocorrelation of $y(t)$	$\vec{r}$	= radius vector in image plane
$\Phi_{xx}(\omega)$	= input spectrum	$\Phi_{ii}(\vec{\rho'})$	= image autocorrelation
$\Phi_{yy}(\omega)$	= output spectrum	$\hat{\upsilon}_{oo}(\vec{\omega})$	= object plane spatial frequency spectrum
$\phi_{yy}(0) = \langle y^2(t)\rangle = \int_{-\infty}^{+\infty} \|H(\omega)\|^2 \Phi_{xx}(\omega)\,d\omega$		$\hat{\upsilon}_{ii}(\vec{\omega})$	= image spectrum
		$\hat{\varsigma}_{ii}(0) = \langle I^2(\vec{r})\rangle = \iint_{-\infty}^{+\infty} \|\tau(\vec{\omega})\|^2 O(\vec{\omega})\,d\vec{\omega}$	
$\phi_{xy}(\tau)$	= input-output correlation	$\phi_{io}(\vec{r})$	= object-image correlation
$\Phi_{xy}(\omega)$	= cross spectrum	$\hat{\upsilon}_{io}(\vec{\omega})$	= cross spectrum

5.6 TREATMENT OF RANDOM WAVEFORMS

For a random signal existing over an infinite interval, the Fourier transform does not exist. Therefore, a statistical description is required. In general, the analytic attack consists of defining the operation of a finite interval. The statistical functions is then the limiting value of the output of the operation as the interval becomes infinitely long. The expressions for Random Input in Table 5-1 stem from this rationale. For example, we have

$$\phi_{xx}(\tau) = \langle X(t)X(t+\tau) \rangle = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} X(t)X(t+\tau)dt$$

(I-5-6)

$$\phi_{oo}(\vec{\rho}) = \langle o(\vec{\rho'})o(\overrightarrow{\rho+\rho'}) \rangle = \lim_{T \to \infty} \frac{1}{4L^2} \int_{-L}^{+L} \int_{-L}^{+L} o(\vec{\rho})o(\overrightarrow{\rho+\rho'})\overrightarrow{d\rho'}$$

This is mathematically correct, however, there is a problem. Since we cannot in actuality average over a record of infinite length, the equation (I-5-6) really instructs us to calculate the mean values over a finite interval T or L. Then we obtain an estimate of the spectrum by taking the Fourier transform of the correlation and letting the interval become indefinitely long, or at least long enough to ensure that the measured S(ω) will meet accuracy requirements. Thus

$$S(\omega) \underset{T \to \infty}{=} S_T(\omega) \underset{T \to \infty}{\rightleftharpoons} \phi_T(\tau) \underset{T \to \infty}{=} \langle X(t)X(t+\tau) \rangle_T$$

(I-5-7)

Unfortunately, the variance of $S_t(\omega)$, in the limit, is very large (Reference 54). Swartz cites Jenkins and Watts (Reference 55) for proof that

$$\lim_{T \to \infty} Var\, S_T(\omega) = S^2(\omega)$$

(I-5-8)

Using the ratio of mean to standard deviation as a measure of signal-to-interference ratio, we find that

$$\frac{E[S_T(\omega)]}{\sqrt{VAR\, S_T(\omega)}} \underset{T \to \infty}{\approx} 1$$

(I-5-9)

This says that the true spectrum is only as big as the noise (uncertainty in $S_t(\omega)$).

The way out of this difficulty is to chop the signal record up into a set of segments. The Fourier transform of each segment is calculated individually to get a set of spectrum estimates, $S_n(\omega)$. The N estimates, $S_n(\omega)$ are then averaged together to get the final estimate of $S(\omega)$.

Now in practice the total observation time, T, which is available is not infinite. The question then arises, "How many segments of duration T/N should we use?" As N gets larger and larger, the variance in the estimate of $S(\omega)$ decreases due to better averaging. However, the resolution of our estimator becomes poorer, because the length of each observation segment is getting shorter. This increases the mean square difference between the true spectrum and estimated spectrum due to loss of resolution. This error is called bias.

It should be evident that there is an optimum N that produces the minimum total error. This total error is noise variance plus bias squared. We will drop this subject for the time being. However, we will see as we progress that preserving a balance between variance and bias is a central theme in analysis and a design driver in mechanization.

The important point at this time is the observation that each signal record is a transient. Therefore, the relations for transient inputs in Table 5-1 can be used. The Random Input relations in the table are no longer mathematical identities, but approximations. Their continued use is important because they eliminate all phase information concerning the successive random signal records. These functions are therefore needed to analyze and design system configurations which must operate in an optimum manner when driven by various members of a class random signals. In contrast to this state of affairs, functions such as x(t) and X(ω) or y(t) and Y(ω) are the signals themselves.

5.7 INTERRELATIONS BETWEEN FUNCTIONS IN TABLE 5-1

It may be seen that there are three classes of functions involved for both time and space filters. One class concerns the input signals. A second function class concerns the output signals. The third class involves the signal transformation process produced by the filters. Stripping away the details of limits, normalization and the like, Tables 5-3, 5-4 and 5-5 show these relations. A star stands for convolution. Dual stars indicate a two-dimensional convolution. Likewise, a dot means multiply. Two dots mean carry out a two-dimensional multiplication. The symbols $\rightleftharpoons$ and $\stackrel{2-d}{\rightleftharpoons}$ indicate one and two dimensional Fourier transforms, respectively. Likewise, single and dual pentagons, $\star$ and $\star\star$ signify single or dual autocorrelations or cross correlations as the case may be.

The relation, $\phi(\tau) \rightleftharpoons \Phi(\omega)$ is an expression of the Weiner-Khinchin theorem. This theorem states that the Fourier transform of the signal autocorrelation function is the power-density spectrum of the signals. Convolutions in the space/time domain are always paralleled by products in the frequency domain as a result of the convolution theorem.

TABLE 5-3. FILTER INPUT RELATIONS

Time Filters		Space Filters	
input amplitude $x(t)$ $\xrightarrow{X(t)\star x(t)}$ **correlation function** $\phi_{xx}(\tau)$		**object intensity** $o(\xi,n)$ $\xrightarrow{o(\zeta,n)\star\star o(\xi,n)}$ **correlation function** $\phi_{oo}(\tau_\xi,\tau_\eta)$	
$\Updownarrow$	$\Updownarrow$ 2-d	$\Updownarrow$ 2-d	$\Updownarrow$ 2-d
$X(\omega)$ $\xrightarrow{X(\omega)\cdot X^*(\omega)}$ $\phi_{xx}(\omega)$		$O(\omega_\xi,\omega_\eta)$ $\xrightarrow{O(\vec{\omega})\cdot\cdot O^*(\vec{\omega})}$ $\phi_{oo}(\omega_\xi,\omega_\eta)$	
amplitude spectrum **power-density spectrum**		**object amplitude spectrum** **power-density spectrum**	

TABLE 5-4. FILTER OUTPUT RELATIONS

Time Filters		Space Filters	
output amplitude $y(t)$ $\xrightarrow{y(t)\star y(t)}$ **correlation function** $\phi_{yy}(t)$		**image intensity** $i(x,y)$ $\xrightarrow{i(x,y)\star\star i(x,y)}$ **correlation function** $\phi_{ii}(\tau_x,\tau_y)$	
$\Updownarrow$	$\Updownarrow$	$\Updownarrow$ 2-d	$\Updownarrow$ 2-d
$Y(\omega)$ $\xrightarrow{Y(\omega)\cdot Y^*(\omega)}$ $\phi_{yy}(\omega)$		$I(\omega_x,\omega_y)$ $\xrightarrow{I(\vec{\omega})\cdot\cdot I^*(\vec{\omega})}$ $\phi_{ii}(\omega_x,\omega_y)$	
amplitude spectrum **power-density spectrum**		**image amplitude spectrum** **power-density spectrum**	

TABLE 5-5. FILTER INPUT/OUTPUT RELATIONS

input amplitude	memory function	output amplitude	object intensity distribution	spread function or antenna power pattern	image intensity distribution
$x(t)$	$\star$ $h(t)$	$=$ $y(t)$	$o(\xi;n)$	$\star\star$ $S(x,y)$	$=$ $i(x,y)$
$\Updownarrow$		$\Updownarrow$	$\Updownarrow$ 2-d		2-d $\Updownarrow$
$X(\omega)$	$\cdot$ $H(\omega)$	$=$ $Y(\omega)$	$O(\omega_\xi,\omega_\eta)$	$\cdot\cdot$ $\tau(\omega_\xi,\omega_\eta)$	$=$ $I(\omega_x,\omega_y)$
input amplitude spectrum	**transfer function**	**output amplitude spectrum**	**object amplitude spectrum**	**modulation transfer function**	**image amplitude spectrum**

5.8 EYE RESPONSE TO THE SPATIAL FREQUENCY SPECTRUM

Just as in the case of temporal filters, one of the great advantages of analysis in the frequency domain is the fact that a series of modulation transfer functions can be multiplied together to obtain the MTF of the total system. A good introduction to this technique is afforded by examination of response of the human eye to the spatial frequency spectrum. This will be followed by a description of the calculation of eye response to a complete photographic system (Reference 56).

Unlike almost all other optical elements, the eye MTF is that of a band-pass filter. Figure 5-3, taken from Price's paper referenced above, shows this response. The low spatial-frequency spectrum components are removed by the retinal neural network which processes the output of the light sensitive elements in the retina (rods and cones). Apparently the reception of the low-frequency components of the scene viewed by the eye had little survival value (or even an adverse effect). These components, therefore, were eliminated by evolutionary processes.

It has been found that optimum photographic systems have a total MTF, including the eye MTF, which has the greatest area under the composite response curve (Reference 56). This condition is subject to various constraints such as constant available light flux and rules for selection of viewing distance. Figure 5-4 is taken from Price's paper. It shows how the total MTF can be found.

The use of this technique in order to generate radiometer images for photointerpretation will be taken up later in this study.

5.9 CONCLUSIONS

The high degree of parallelism between temporal and spatial filters would indicate that the concept of spatial frequencies and spatial frequency spectrums should be very important when designing lens systems or antenna systems. This indeed is true in optics. The modulation transfer functions of lenses and other optical components has become the standard method of specifying their performance. Design of lenses by use of frequency-domain concepts have become the dominant design methods. Recursive algorithms for computer lens design use very sophisticated descriptions of the response of the human eye to determine what changes in the lens design should be made on each recursion (Reference 56).

Radio-astronomers have made extensive use of spatial-frequency concepts in order to design radio telescopes. Unfortunately, most antenna engineers have not followed suit. Often this has led to designs which appear primitive compared to those of optical engineers and radio astronomers. This study will concern itself almost entirely with these frequency-domain design techniques as a means of designing radiometric mappers. This concluding section seems to be a good place to mention one other interesting correspondence between spatial and temporal frequencies. In addition, some differences between space and time filters will be mentioned which are not apparent from the tables.
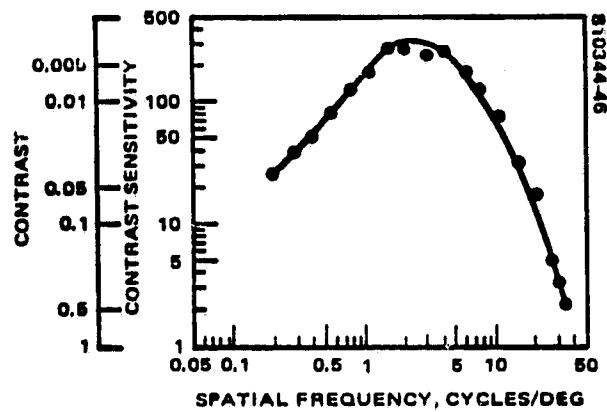
FIGURE 5-3. MODULATION TRANSFER
FUNCTION OF THE HUMAN EYE
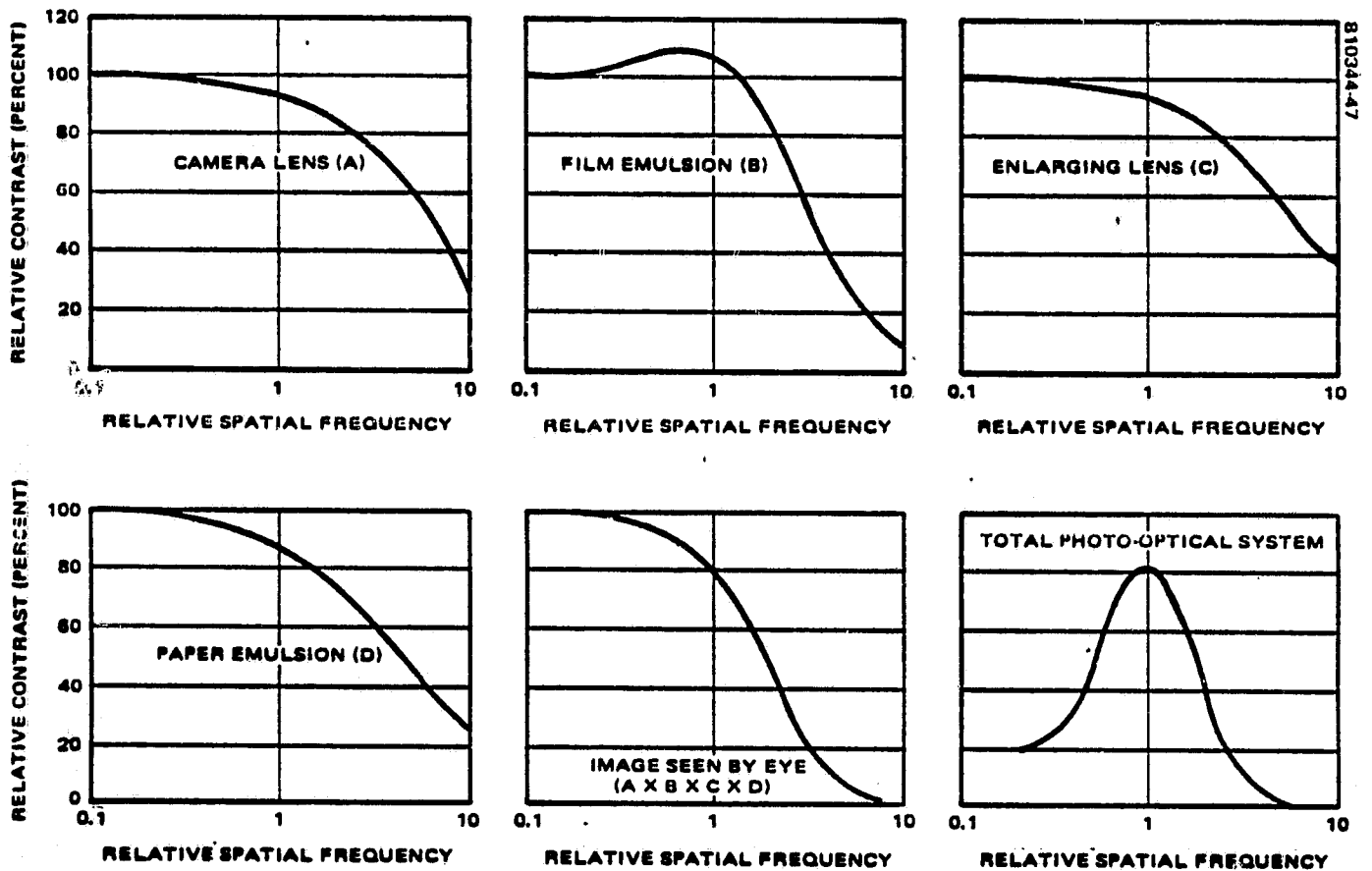
FIGURE 5-4. COMPOSITE MODULATION TRANSFER FUNCTION

In optical/antenna systems light distributions fall on both sides of the coordinate axes. Indeed, these distributions often are symmetric around these axes. This means that the causality condition, which exists when t is the independent variable, does not exist in the spatial frequency domain. Therefore, the real and imaginary parts of the transfer function, $\tau(\omega_\xi, \omega_\eta)$, are not a Hilbert transform pair as in the temporal case.

Also, in general, in the temporal case frequency is a complex variable. That is $S = \sigma + i\omega$. Then h(t) and H(ω) are related by a generalization of the Fourier transform called the Laplace transform. Spatial frequencies are always real, so the Laplace transform is never required.

Neither of these differences harm the usefulness of the correspondence between temporal and spatial frequency functions in aiding the analysis and design of optical and antenna systems.

There is an interesting correspondence between the uncertainty relations in both domains. We all know that for temporal frequencies the uncertainty relation is

$$\frac{1}{2\pi} \Delta\omega\Delta\tau \geq 1 \tag{I-5-10}$$

The corresponding relation for spatial frequencies is (Reference 57)

$$\frac{1}{2\pi} \Delta\rho\Delta x \geq 1 \tag{I-5-11}$$

where Δx is the size of a pixel resolved by the optical spatial channel and $\Delta\rho$ is the spatial bandwidth in radians per unit distance.

6. ANTENNAS AS SPATIAL FREQUENCY FILTERS

6.1 RELATION OF ANTENNA MTF TO APERTURE EXCITATION

In Section 5 the concept of the modulation transfer function (MTF) of a lens or antenna was introduced. It was shown in Table 5-5 "Input-Output Relations" as the Fourier transform of the filter memory function in the temporal case. It also was shown as the Fourier transform of a spread function in the space-filter case. If the device is an antenna, the spread function is called the antenna power pattern, $P(\phi)$.

It has been shown by Booker and Clemmow (Reference 59) that the antenna power pattern is proportional to the complex autocorrelation function of the aperture distribution function. Thus

$$\tau(x_{\lambda_0}) \propto E(x_\lambda) \star E^*(x_\lambda) \qquad \text{(I-6-1)}$$

where $\tau(x_{\lambda_0})$ is Fourier transform of the antenna power pattern, $E(x_\lambda)$ equals the aperture distribution, $x_\lambda = x/\lambda =$ distance across the aperture face in wavelengths.

One would have expected Eq. (I-6-1) to be a convolution, rather than autocorrelation. Kraus (Reference 59) shows that when Eq. (I-6-1) is used, the antenna response to a brightness distribution function is

$$S(\phi_0) = B(\phi) * \tilde{P}(\phi) \qquad \text{(I-6-2)}$$

where $\tilde{P}(\phi)$ is a mirror image of the power pattern $P(\phi)$. This is equivalent to the space filter relation in Table 5-5. However, an angular response is found, instead of a response over an image plane. This mirror image pattern is used so that the characteristic folding and inversion operation required by convolution will not be needed in Eq. (I-6-1). The autocorrelation replacing the convolution is more difficult. Table 6-1 shows two tableau. The first shows where Eq. (I-6-1) comes from. The second justifies the corresponding convolutional form of Eq. (I-6-1). The hat symbol $\sim$ indicates a mirror-image function formed by inversion; e.g., $\tilde{P}(\phi_0 - \phi) = P(\phi - \phi_0)$. The variable x_{λ_0} in Eq. (I-6-1) is a spatial frequency. It will be replaced by the symbol, F.

TABLE 6-1. CALCULATION OF THE MODULATION TRANSFER FUNCTION
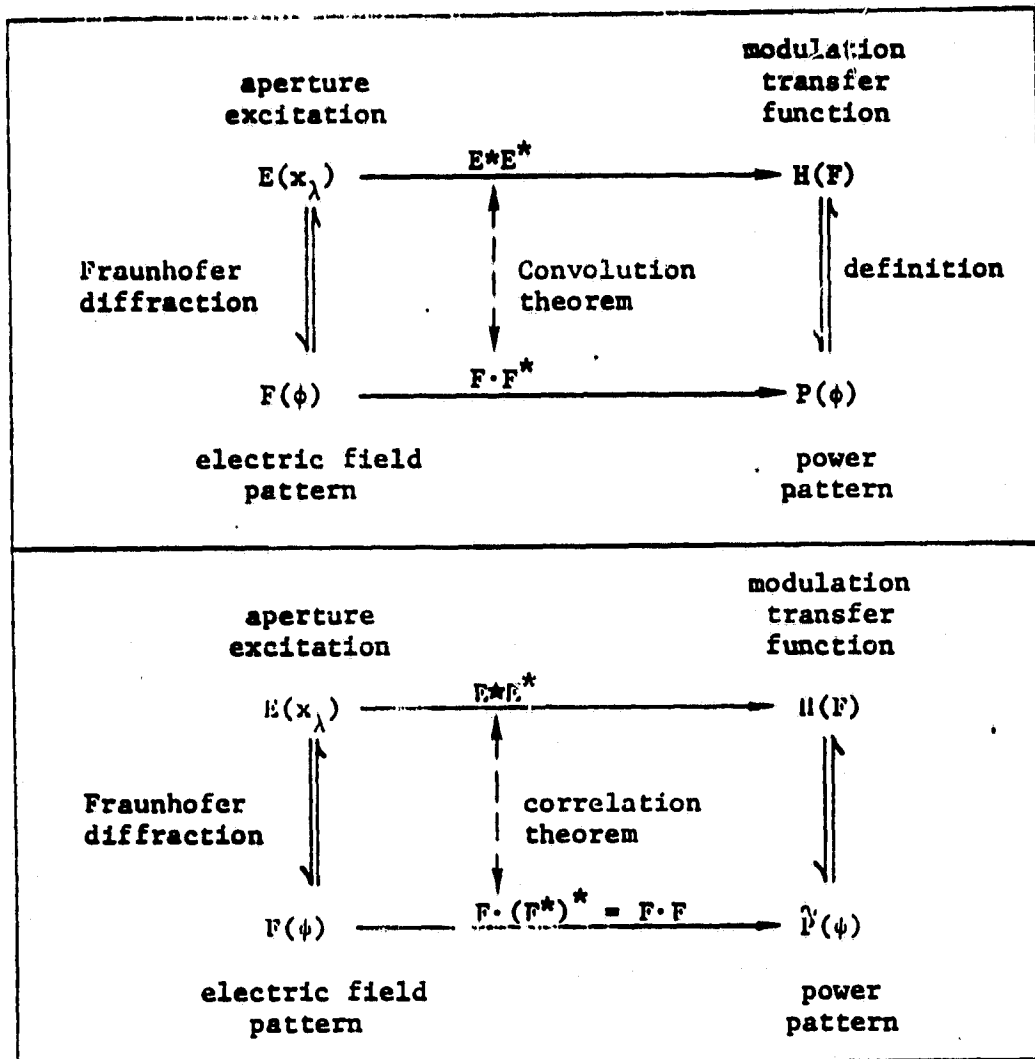
The correlation theorem is not as well known as the convolution theorem. It is discussed by Brigham (Reference 60). If E is symmetric, the distribution between autocorrelation and convolution disappears.

6.2 ANTENNA MTF FOR SOME SPECIAL EXCITATION FUNCTIONS

Figure 6-1 shows a graphic method of finding the MTF when the excitation is given. All autocorrelations slope off to zero in manner shown. The equivalent of the general tableau shown in Table 6-1 is given in Figure 6-2 for uniform weighting.

X = APERTURE
COORDINATE,
WAVELENGTHS

$a_\lambda = a_\lambda$ = APERTURE
SIZE,
WAVELENGTHS

a_λ = CUTOFF SPATIAL FREQ

ΔU_{NN} = $2/a_\lambda$ = NULL-TO-NULL BEAMWIDTH

$1/a_\lambda$ = ANGLE SUBTENDED BY 1 CYCLE OF
CUTOFF SPATIAL FREQ
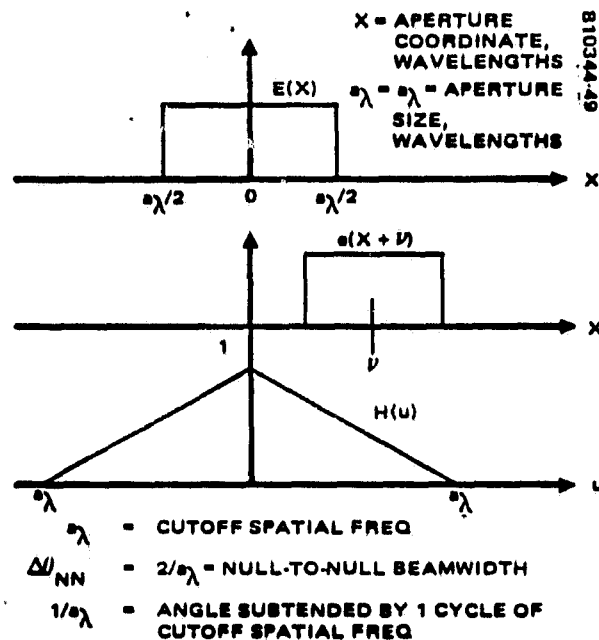
FIGURE 6-1. GRAPHIC CALCULATION OF
A MTF

RECT (X_λ) —— RECT + RECT* —— TRI(F)

SINC (Q) —— SINC • SINC* —— $SINC^2(Q)$
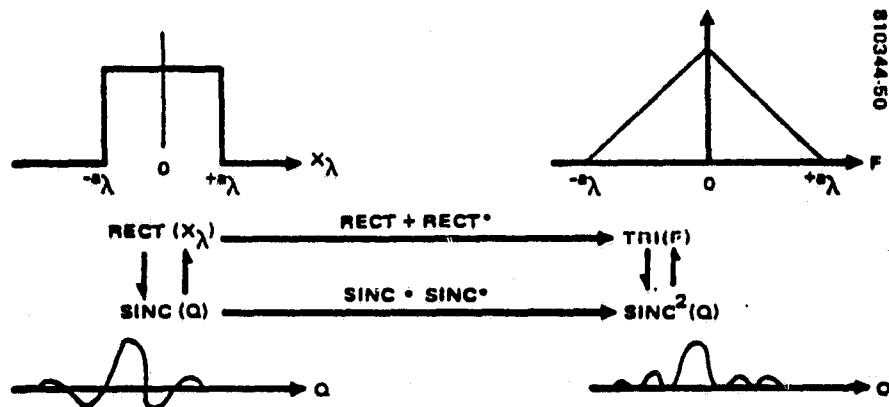
FIGURE 6-2. FUNCTIONS ASSOCIATED WITH UNIFORM
APERTURE EXCITATION

Note that TRI(F) reaches zero at $\pm$a. This is called the cutoff spatial
frequency. The MTF is zero outside this interval. As noted, the null-to-null
beamwidth of the power pattern, $sinc^2\phi$, is $2/a\lambda$ radians. Since the period
of the cutoff frequency is $1/a\lambda$, the cutoff wavelength executes two cycles
per null-to-null beamwidth. The Nyquist criterion then says that if the
antenna beam is scanned, at least four samples of the antenna output must be
taken during the time the beam is scanned through an angle equal to the null-
to-null beamwidth of the beam.

It can be seen from Figure 6-1 that the cutoff spatial frequency will always be equal to the aperture size in wavelengths, a_λ. As an example of a more complicated excitation function, let $E(x_\lambda) = TRI(x_\lambda)$. We have found in Figure 6-2 that:

$$TRI(y) \Longrightarrow sinc^2(z).$$

So the tableau for this case must be as shown in Figure 6-3. Tables of modulation transfer functions for more complicated types of aperture excitation weighting have been tabulated and are available (Reference 62).

All filled apertures act like bandpass filters with response which goes to zero at $\pm a_\lambda$. It is obvious that in order to pass the spatial frequency spectrum with no distortion from zero out to the spatial frequency a_λ, the antenna MTF should be rectangular. If the total spectrum of the scene fell in this 0 to a_λ range, there would be no signal distortion. In the more likely case where the spectrum extends beyond a_λ the rectangular MTF produces the least distortion possible without increasing system resolution. Bracewell calls the resulting image signal the principal solution.

At first blush it would appear that regardless of what aperture weighting is chosen, the MTF of an aperture with continuous excitation must slope downward until a value of zero is reached at a_λ. Probably the closest approach to a rectangular response that can be obtained is produced by use of Taylor weighting with fairly high sidelobes. For instance, the MTF of Taylor weighted antenna with -5 dB sidelobes shown in Figure 6-4 comes as close to a rectangular response as is possible (Reference 62).
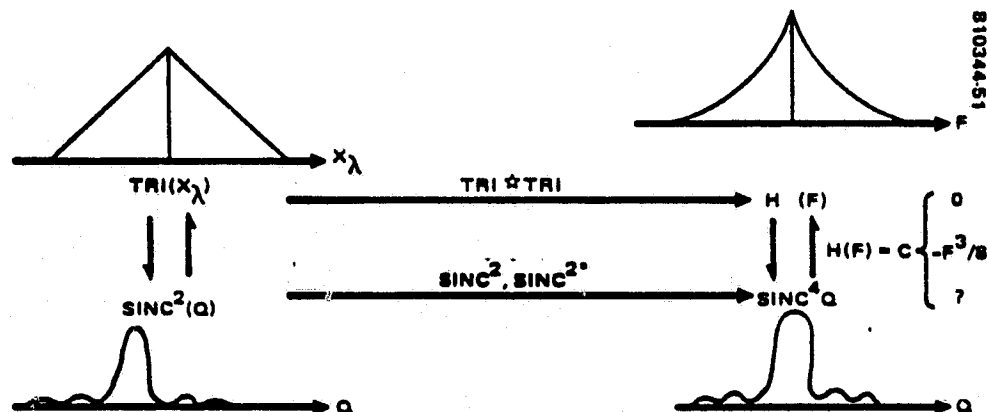
FIGURE 6-3. FUNCTIONS ASSOCIATED WITH TRIANGULAR APERTURE EXCITATION
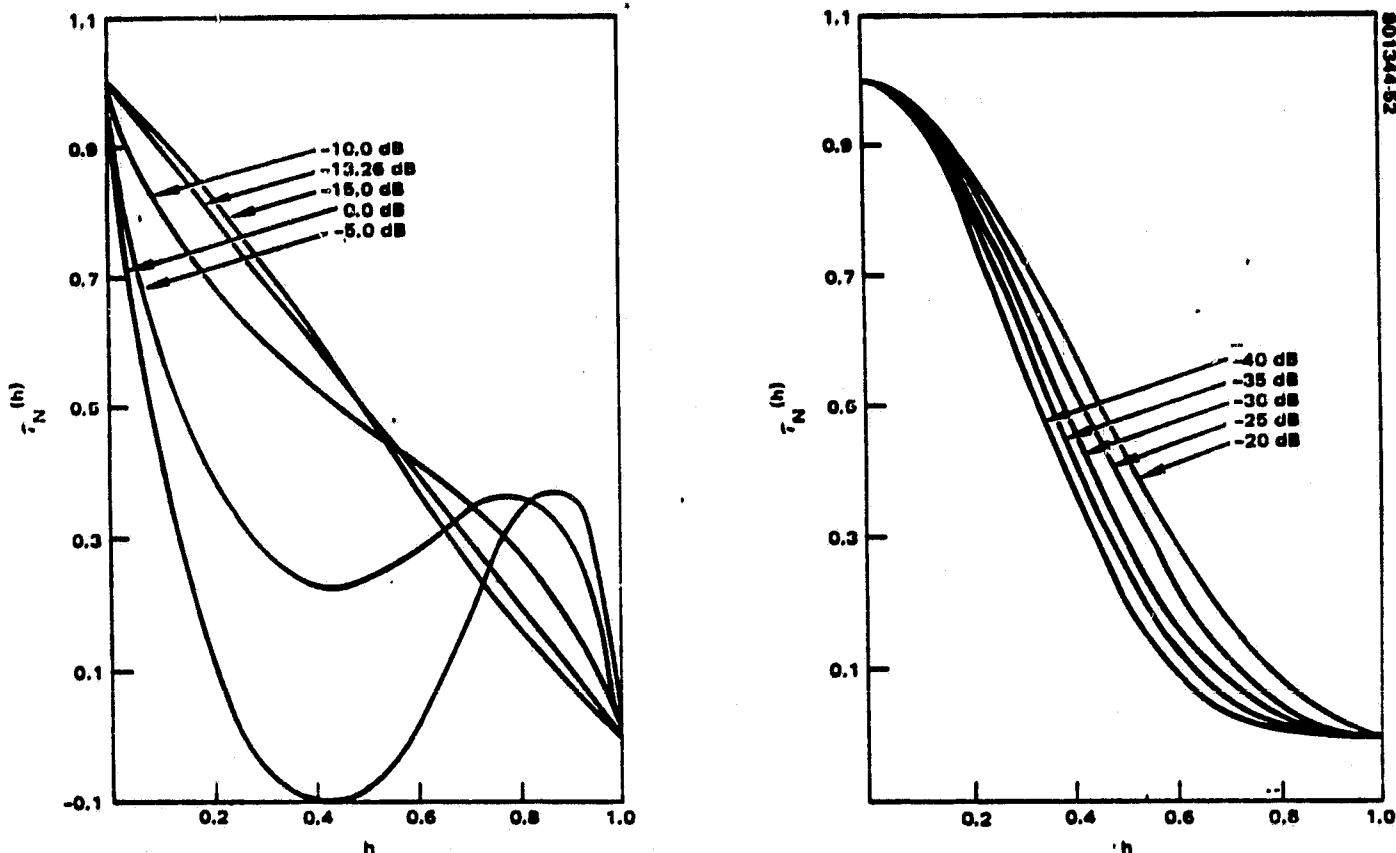
FIGURE 6-4. MODULATION TRANSFER FUNCTIONS OF TAYLOR WEIGHTED APERTURES

6.3 MTF OF THE COINCIDENT ARRAY OF UTUKURI AND McPHIE

The trouble lies in the power pattern of an antenna with a rectangular MTF. This principal power pattern is the Fourier transform of the rectangular MTF. This power pattern is a sinc ϕ response. It has both positive and negative sidelobes. However, a radiometer using a square-law detector cannot be used to obtain the princiapl power pattern because the detector output cannot be negative.

However Utukuri and McPhie get around this difficulty by the use of two interleaved array structures. Instead of using a square-law detector, the two array outputs are multiplied together. This arrangement is shown in Figure 6-5 (Reference 63).

It can be seen that Utukuri and McPhie have replaced the power-measuring radiometer associated with a single antenna with a correlating radiometer. The two radiometer channels are fed by the pair of interleaved arrays.

Figure 6-6 shows the calculated response of a five-element array. The sidelobes of alternating polarity may be seen.
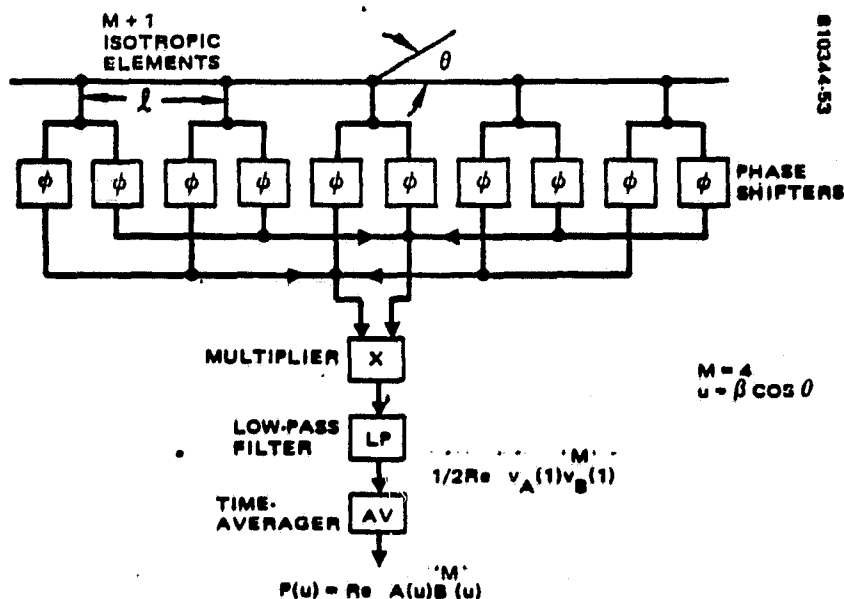
FIGURE 6-5. COINCIDENT ARRAY OF UTUKURI AND MACPHIE
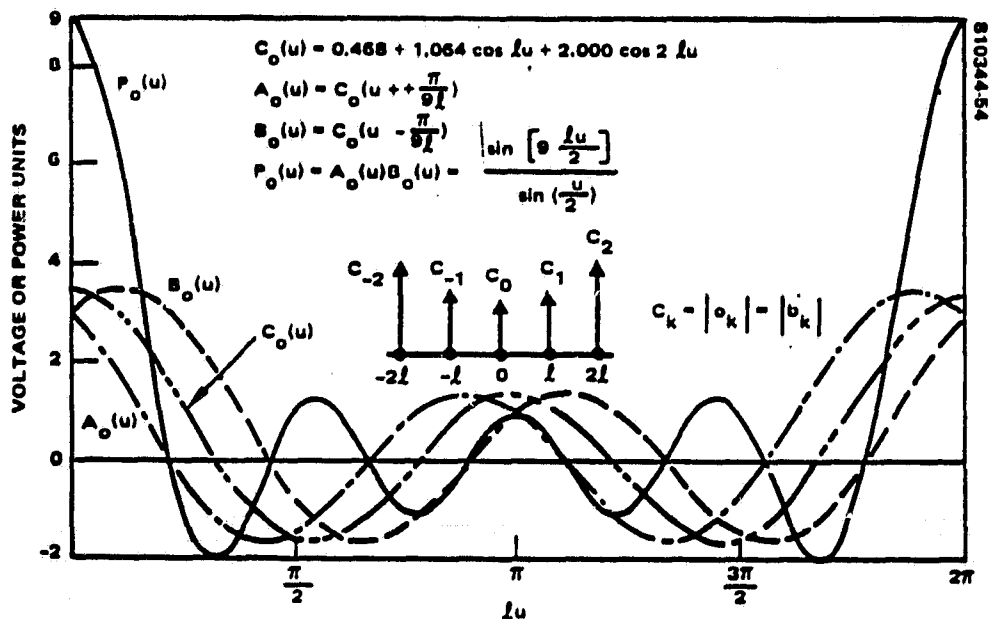
FIGURE 6-6. FACTOR PATTERNS AND PRINCIPAL POWER PATTERN
OF A FIVE ELEMENT COINCIDENT ARRAY

6.4 MTF OF A MULTIPLYING INTERFEROMETER

We are now ready to go on to the graphic calculation of the modulation
transfer function of a multiplying interferometer. Figure 6-7 showing this
graphic solution is due to Kraus (Reference 59), and is his Figure 6-22.
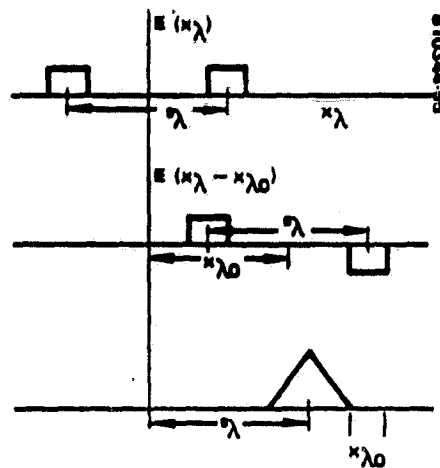
FIGURE 6-7. GRAPHIC DETERMINATION
OF THE MODULATION TRANSFER OF A
MULTIPLYING INTERFEROMETER

Thus we see that the MTF of a multiplying interferometer has the spatial frequency response of a bandpass filter.

Kraus in his Eq. (6-100), pg 184 gives the interferometer power pattern function for the case where the antenna elements at the ends of the interferometer are dissimilar. It is

$$P(\phi) = E_1(\phi)E_2^*(\phi)e^{+j\psi} + E_1^*(\phi)E_2(\phi)e^{-j\psi} \tag{I-6-3}$$

This result is proved in Appendix B "Generalized Interferometer Mapping Behavior." Eq. (I-6-3) of this section is given in Eqs. (B-15) and B-17) of Appendix B in terms of the gain functions of the two antenna elements with their phase centers moved to the coordinate origin at the midpoint of the baseline. The functions $E_1(\phi)$ and $E_2(\phi)$ are the field patterns of the two interferometer elements and $\psi = 2\pi S_\lambda \sin\phi$.

If the elements are identical, that is $E_1(\phi) = E_2(\phi) = E_0(\phi)$, then the power pattern becomes

$$P(\phi) = |E_0(\phi)|^2 \cos(2\pi S_\lambda \sin\phi) \tag{I-6-4}$$

Finally, if the element patterns are symmetrical, so that $E_1(\phi) = E_1^*(\phi)$ and $E_2(\phi) = E_2^*(\phi)$, the relative power pattern becomes

$$P(\phi) = E_1(\phi)E_2(\phi)\cos\psi \tag{I-6-5}$$

The power pattern is seen to be proportional to the product of the field patterns of the individual elements.

6.5 METHODS OF VARYING THE EFFECTIVE BASELINE LENGTH

These mapping radiometers must vary the fringe pattern given by Eq. (I-6-4) in order to carry out their mapping functions. This can be done in a variety of ways. One method of changing the fringe pattern is by changing the baseline length, S_λ, by moving the antenna elements. We can also change S_λ for a fixed antenna separation by sweeping the frequency. Both of these methods change the spacing between lobe peaks, as well as lobe position.

If a phase shifter is inserted in series with one element, ψ becomes $(2\pi S_\lambda \sin\phi + \phi_S)$. If ϕ_S is continually increased by the phase shifter, a source at fixed position will produce a sinusoidal voltage at the interferometer output. Alternately we can put a variable length transmission line in series with one of the elements. Then

$$\psi = 2\pi\frac{\nu}{c}(Ss\text{in}\phi + \Delta\ell/p) \qquad\qquad (I-6-6)$$

Where $\Delta\ell$ = differential length of the transmission line feeds from the radiometer, p = transmission line phase velocity = ν/c, ν = velocity of a signal along the line.

The last two methods are equivalent if $\Delta\ell$ is small. However, they behave entirely differently when $\Delta\ell$ is large.

If the interferometer elements were points, then the bandpass filter response would become an impulse. However, if the elements are not points, then the MTF passband has a finite width as shown in Figure 6-7. This will produce a bias error when measuring the magnitude of the spatial frequency spectrum at frequency S_λ. Correction of this bias error will be considered in a later section.

7. FREQUENCY DOMAIN IMAGERS

7.1 FOURIER OPTICAL SYSTEMS

Fourier optics has become an important subject in recent years. (References 64, 65) These techniques involve conversion of optical images to spatial frequency spectrums, recording such spectrums in the form of photographic records called holograms, and filtering these spectrum records. The results of these operations are used to reconstruct modified images by transformation back to the space domain.

Usually the objects dealt with in optics are recorded as images on photographic film. Manipulation of the image starts with the photo record. In our case, the input signal is in the form of microwaves radiated by a scene at a great distance from the radiometer. We want to record data about the radiation over a plane corresponding to one surface of an imaging lens as shown in Figure 7-1.

In the optics case the image on the photo transparency is converted into a spectrum by use of a lens. Use is made of the fact that lenses in the correct optical configuration can act as two-dimensional Fourier transform devices. The transformed image appears as a spectrum in a plane which is called a frequency plane or Fourier plane.
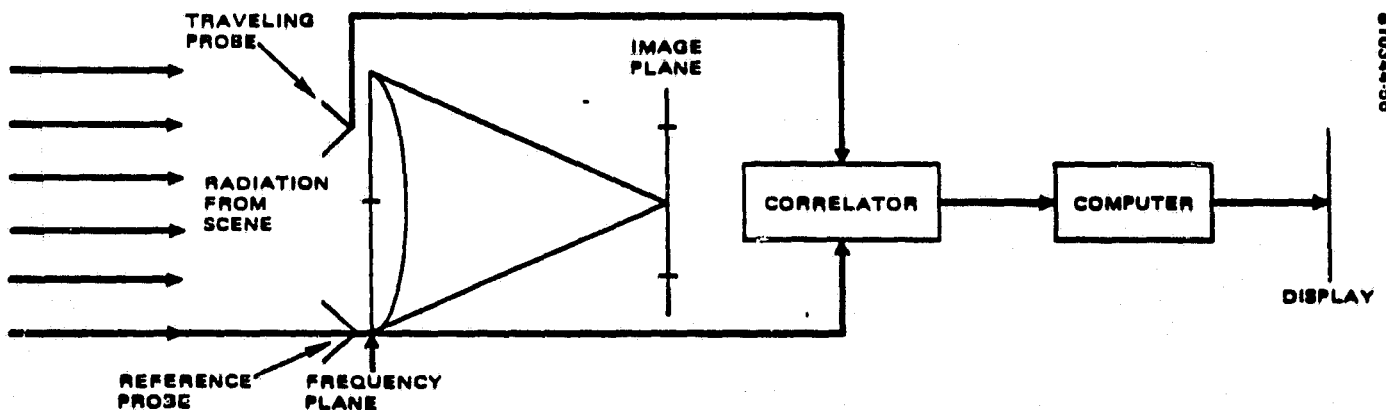
FIGURE 7-1. DIRECT IMAGING SYSTEM USING A LENS CONTRASTED WITH FREQUENCY-DOMAIN IMAGER USING TWO RADIATION-FIELD PROBES

In our case the distant scene produces a frequency-domain field at the surface of the imaging lens. This surface, shown in Figure 7-1, is a frequency plane. The lens acts as a Fourier transformer which converts the data to the space domain. This converted signal is the image which appears in the image plane.

We intend to record the frequency-plane data. Then the image can be generated by carrying out the Fourier transform by data processing in a computer. The question then arises as to exactly what format should be used to record the frequency plane data. SAR radars give some guidance on this point.

7.2 SYNTHETIC-APERTURE RADARS AS ACTIVE FREQUENCY-DOMAIN IMAGERS

SAR radars are active frequency-domain imagers. The coherent video, which is recorded as the synthetic aperture is generated, gives us a clue to required format in the passive case. The SAR antenna receives the radiation reflected by the scene. The phase change from pulse to pulse of the signal return is the most important part of the data about the radar return. In order to tie the data from successive returns into a single unbroken record of the phase changes there must be a master timer signal. This signal is generated by very stable crystal oscillators. It is called the reference signal. Its smooth, predictable phase change is compared with the unknown phase change of the return by means of a phase detector (synchronous demodulator). This is shown in Figure 7-2.
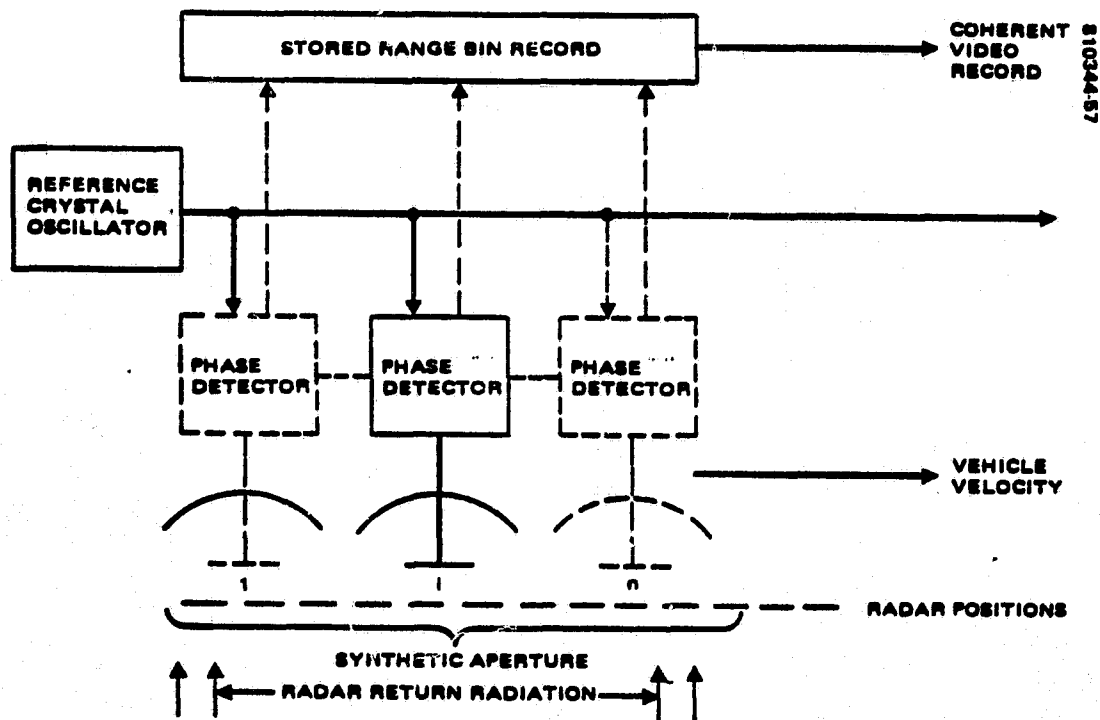
FIGURE 7-2. SAR MEASUREMENT IN THE FREQUENCY PLANE

The coherent video record is Fourier transformed to get the image or Fresnel transformed when the scene is in the near-field of the synthetic aperture (the Fresnel transform is a generalization of the Fourier transform). The radar signal is linear in amplitude, while the radiometer signal is linear in power. However, this is not important in making this comparison. The mathematics is the same.

7.3   THE VAN CITTERT-ZERNIKE THEOREM

The radiometer imager in Figure 7-1 has no internal reference like the SAR. However, all we need is a reference with a stable phase derived from the radiation from the scene.

This can be provided by a second probe with a fixed location in the frequency plane. The fact that the probe receives noise, rather than a sine wave, is not important. As long as the noise bandwidth is small compared to the center frequency of the noise spectrum a phase comparison can be made. The generalization of the phase detector, when noise-like signals are used, is the correlator. There the outputs of the reference and traveling probes are multiplied together and averaged.

We readily will recognize the output of the correlator as a lateral coherence function, Γ_{12} (τ_ξ, τ_η). Its importance stems from the fact that it is the Fourier transform of the object intensity distribution shown in Table 5-3. This relation is the Van Cittert-Zernike Theorem. This theorem is the basis for the analysis of the propagation of partially coherent radiation. A proof of this theorem, abstracted from Born and Wolf (Reference 66), is given as Appendix E.

This theorem says that if the linear dimensions of the source and the distance between the two frequency-plane probes are small compared to the distance between the frequency plane and the source, the degree of coherency between these probes, Γ_{12} (τ_ξ, τ_η), is equal to the absolute value of the Fourier transform of the intensity function of the source.

In Table 5-1 relations between functions in the object plane, in the image plane, and relations between these function sets were given. We now need to prepare a similar tableau for object-plane/frequency-plane relations. This can be done with the aid of the Van Cittert-Zernike theorem. This is done in Figure 7-3. The symbol ← α → indicates that the two functions differ only by a constant factor.

VAN-CITTERT ZERNIKE
THEOREM

OBJECT
INTENSITY
DISTRIBUTION

$0 (\xi, \eta)$ ⟷ $\Gamma_{12} (\tau\xi, \tau\eta)$

2-d

LATERAL
COHERENCE
FUNCTION

TRANSFORM
PAIR
DESCRIBING
SOURCE

2-d

a

a

FUNCTION PAIR
DESCRIBING
FREQUENCY
PLANE

OBJECT
AMPLITUDE
SPECTRUM

$0 (\omega\xi, \omega\eta)$ ⟶ $S(V, \nu)$
$\omega\xi$ AND $\omega\eta$ ARE
IN CYCLES
PER UNIT LENGTH

V AND ν ARE
IN CYCLES
PER RADIAN

SPATIAL
FREQUENCY
SPECTRUM

FIGURE 7-3. SOURCE/FREQUENCY PLANE RELATIONS
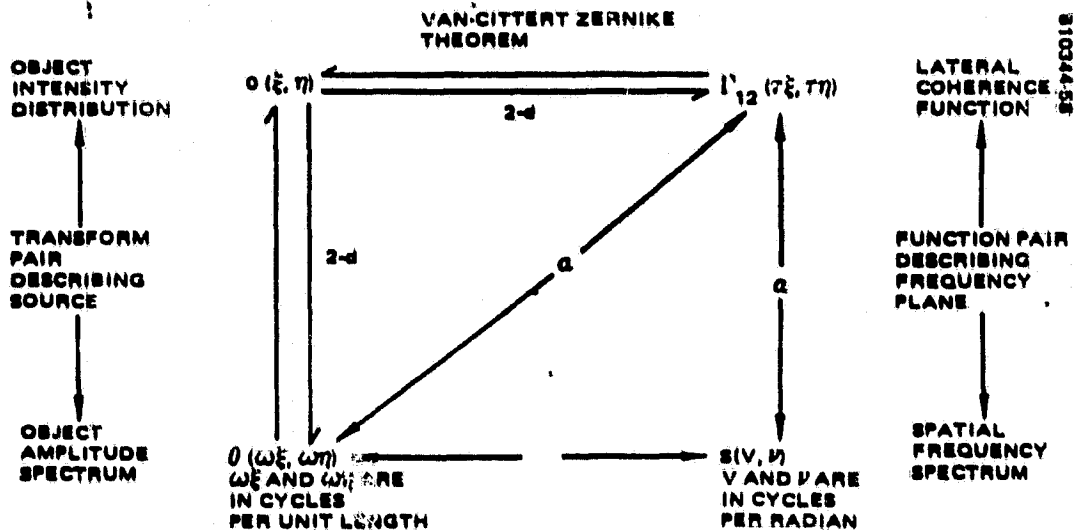
7.4 DUAL INTERPRETATION OF FREQUENCY-PLANE INTERFEROMETER MEASUREMENTS

Since $0 (\xi, \eta)$ is the transform of Γ_{12} and $\vartheta (\omega_\xi, \omega_\eta)$ is a transform of $0 (\xi, \eta)$, $\vartheta (\omega_\xi, \omega_\eta)$ must be proportional to Γ_{12}. The spatial frequency spectrum S (υ, ν) is not the transform of Γ_{12} because it is not a spectrum description of the variations of Γ_{12} as τ_ξ and τ_η are varied. As shown in Section V, S (υ, ν) is the angular spectrum of the incoming radiation. In Figure 7-4 a geometric construction shows the $\vartheta (\omega_\xi, \omega_\eta)$ and S (υ, ν) are proportional. Therefore, $\Gamma_{12} (\tau_\xi, \tau_\eta)$ and S (υ, ν) are proportional.

$\phi_{(\tau)}$

CORRELATOR

BASELINE LENGTH   EQUALS Z

+   +   -   -   +   +

$0 (\omega_\xi, 0)$

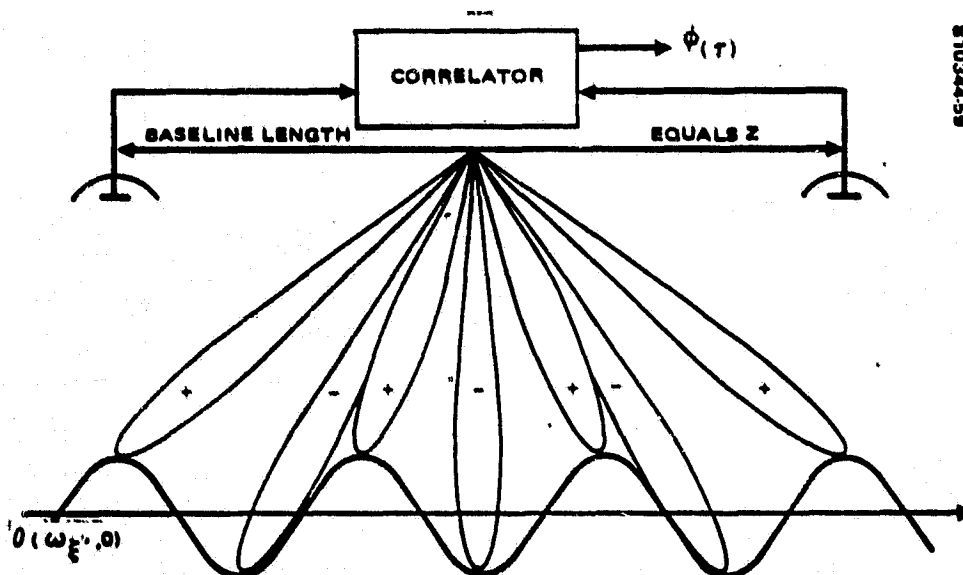FIGURE 7-4. EQUIVALENCE OF ϑ AND S FUNCTIONS

Figure 7-4 shows the two probe antennas separated by a baseline of length Z. This baseline length produces a set of lobes with angular frequency of v^1 (cycles per radian). There will be a component of the scene brightness which fluctuates at a linear frequency of ω_E^1 cycles per unit length. For an interferometer distance from the scene of $\overline{OF}$; the interferometer lobes are synchronous with the brightness component of frequency ω_E^1. A large output of magnitude S (v^1, o) or ϑ (ω_E^1, o) appears at the interferometer output port. All other components of the object brightness distribution average to zero at the output port. This occurs because the interferometer lobes are not synchronous with any of these other spectral components.

We see then that the correlation function measured by the frequency plane probes has a dual interpretation. It is either the lateral coherence function of the frequency plane or the spatial frequency spectrum of the scene. In either case Z must be varied through a range of values in order to measure the total function. These arguments prove the validity of the bottom and right-hand proportionalities in Figure 7-3.

It is now obvious that by moving a probe over a region in the frequency plane, and measuring Γ_{12} as we move, we are in a position to find the object intensity distribution. This is done by taking the two-dimensional Fourier transform of the measured lateral coherence function, Γ_{12} (τ_ξ, τ_η). This process is called aperture synthesis. The probe/correlator device, plus the Fourier transformer, constitutes a frequency-domain imager.

7.5 RECIPROCAL RELATION BETWEEN SPACE-AND FREQUENCY-DOMAIN IMAGERS

Transform pairs always exhibit a reciprocity in width. As one member of the pair becomes smaller the other member becomes larger and vice versa. The reciprocal width theorem states this property more exactly. It says that the equivalent width of a function is equal to the reciprocal of the equivalent width of its transform (Reference 67).

If the frequency-domain imager is contrasted with a pencil-beam scanner (space-domain imager), there are two transform pairs whose comparison is of interest. One pair is probe travel required to produce the image. The other pair is antenna size. We will make the comparison using a microwave lens in the pencil-beam scanner case. This is not a restriction on the relations involved. The lens merely eliminates the folded ray paths produced by reflector antennas. This makes for a neater diagram. Figure 7-5 shows the two cases.

It is Ko (Reference 68) who gives the relation $D = \Delta\theta^{-1}$, where D is the diameter of the disk sampled in the Fourier plane in wavelengths and $\Delta\theta$ is angular resolution. He also states that in order to meet the Nyquist sample criterion, so there is no aliasing, the sample spacing in wavelengths must be $\overline{FOV}^{-1}$. However, in order for the probe antenna to see the entire scene, it must have an aperture of no greater than $\overline{FOV}^{-1}$ wavelengths. Therefore the probe moves by its aperture width, between successive samples.
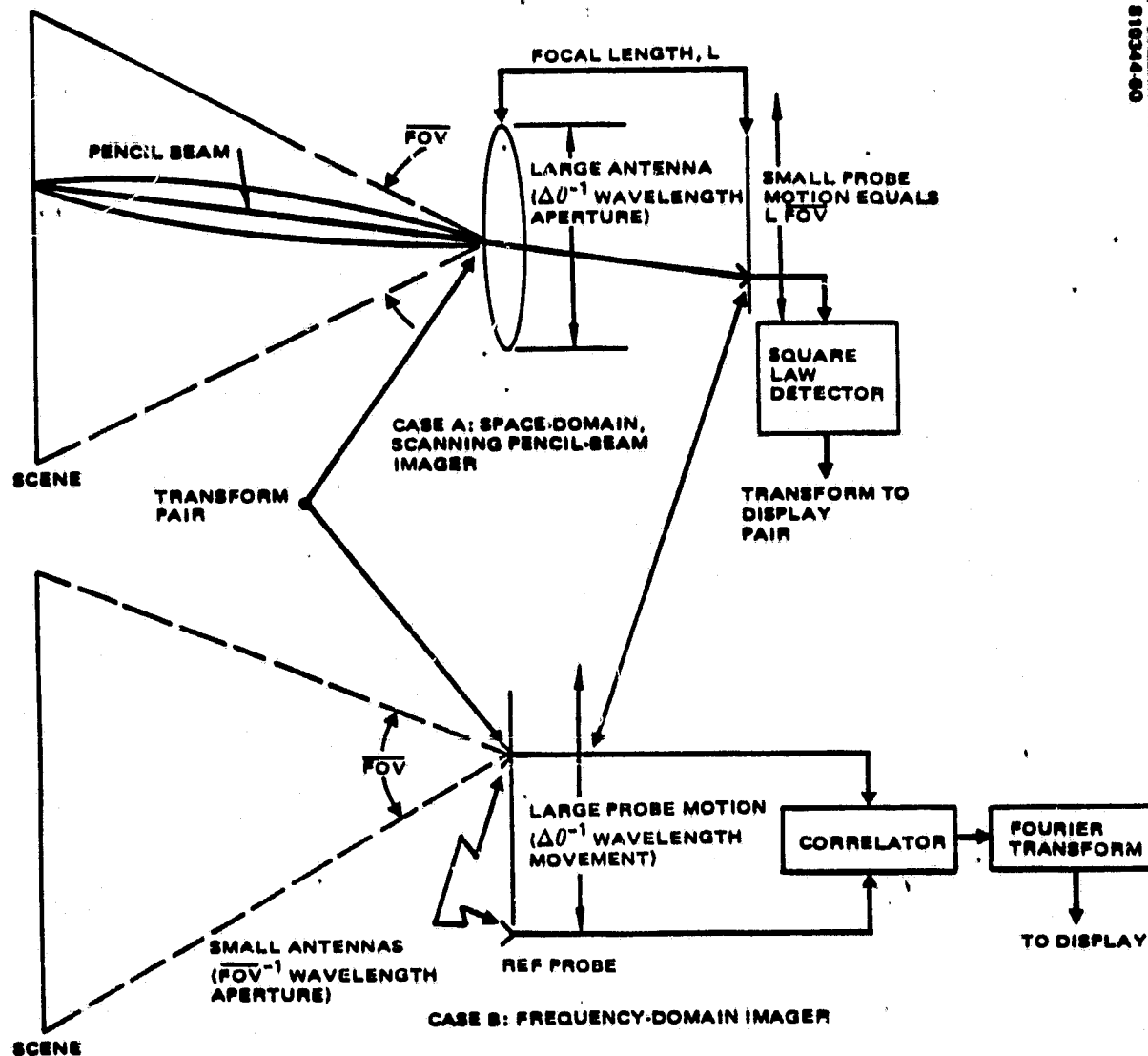
FIGURE 7-5. COMPARISON OF SPACE AND FREQUENCY-DOMAIN IMAGERS

It may be seen that the principle of transform reciprocity favors the frequency-domain imager. This device requires a small antenna and large probe motion. A mapper where only large motions are needed to obtain high resolution is ideal for space applications. There is a lot of real estate in orbit. Motion over the frequency plane is limited only by the ingenuity of antenna engineer and the vehicle designer.

7.6 COMPARISON OF THE SENSITIVITY OF SPACE-AND FREQUENCY-DOMAIN IMAGERS

In the Introduction heuristic arguments were given to prove that the sensitivities of these transform twins were the same. The time has come to prove this on a more formal basis. Rather than dealing with brightness, as we have up to now, let us consider the total brightness of the scene when viewed at the location of the radiometer antenna. Brightness is a spectral density. Total brightness, $B(\theta, \phi)$ is the integral of the brightness over the radiometer operating band. It has the dimensions of watts $m^{-2} rad^{-2}$. Since we are not dealing with a particular scene, we will replace $B(\theta, \phi)$ with an average total brightness, $\bar{B}$, over the field-of-view of the antennas.

The antenna has a capture cross section $A(\theta, \phi)$. The incremental power received from direction (θ, ϕ) is $B(\theta, \phi) A(\theta, \phi) d\Omega$, where $d\Omega$ is a differential solid angle. This cross section can be related to the antenna power pattern by the equation

$$A(\theta, \phi) = A_e P_N(\theta, \phi) \tag{I-7-1}$$

This normalized power pattern is dimensionless. Its peak value is unity. A_e is the antenna effective aperture for radiation angle-of-arrival in the direction of peak gain.

A very important property of antennas is the fact that the integral of $A(\theta, \phi)$ over a sphere surrounding any antenna is always λ^2 (Reference 69). This is true regardless of the antenna size.

Let us assume that the beam efficiencies of both the pencil-beam scanning antenna and the frequency-plane probe are 100 percent. Then in the first case the integral of $A(\theta, \phi)$ over the pencil-beam will be λ^2. In the second case the integral of $A(\theta, \phi)$ over the field-of-view seen by the frequency-plane probe will be λ^2. The solid angles corresponding to these cases will be designated Ω_1 and Ω_2.

The brightness of any emitting surface is given by the relation

$$B(\theta, \phi) = \frac{2kT(\theta, \phi) \Delta\nu}{\lambda^2} \varepsilon(\theta, \phi) \tag{I-7-2}$$

where T is the surface temperature and ϵ is the emissivity. The average brightness is given by

$$\overline{B} = \frac{2\kappa\Delta\nu}{A\lambda^2} \int\int_{scene} T(\theta, \phi)\epsilon dA = \frac{2\kappa\overline{T}\Delta\nu}{\lambda^2} \qquad (I-7-3)$$

where $\overline{T}$ is the average radiant temperature of the scene.

Then for both the pencil-beam antenna and frequency-plane probe cases the thermal power received is

$$P_{1\ or\ 2} = \frac{1}{2} \int\int_{\Omega_1\ or\ \Omega_2} B(\theta, \phi) A(\theta, \phi) d\Omega = \frac{\overline{B}}{2} \int\int_{\Omega_1\ or\ \Omega_2} A(\theta, \phi) d\Omega$$

$$\qquad (I-7-4)$$

$$= \frac{1}{2} \left(\frac{2\kappa\overline{T}\Delta\nu}{\lambda^2} \right) \lambda^2 = \kappa\overline{T}\Delta\nu$$

The factor 1/2 appears because the antennas can receive only one of two orthogonal polarization components of the radiation.

It may be seen that the large pencil-beam, scanning antenna and the small frequency-plane probe receive the same power. In actuality the frequency-plane imager receives twice the power found in Eq. (I-7-4). This occurs because there are two frequency-plane probes. However, the number of samples in the Fourier plane is double the number of pixels in the image scanned by the pencil beam. Call this number N. Then the 2N spectrum samples can be N complex numbers. The total power distributed between the real and imaginary parts of the complex number is P_2 in Eq. (I-7-4). However both real and imaginary bins contain the same noise power as that in a pixel bin. This results because there is the same time available for measurement of the N samples or N pixels. Therefore the sample noise power is twice the pixel noise power. But the two Fourier-plane probes receive twice the signal power compared to the power received from a pixel.

The final conclusion is that the space-domain and frequency-domain imagers have equal sensitivities when all system parameters are the same in both cases.

7.7 SAMPLING THE FREQUENCY PLANE

Up to this point the sampling procedure followed by the traveling frequency-plane probe has been vague. It has only been specified that the sample grid be dense enough to meet the Nyquist criterion. Usually the sampling sequence is divided into two steps. The first step consists of moving the traveling probe out along a baseline, starting at the reference probe and terminating at some maximum baseline length. A new baseline orientation is then selected by stepping the baseline drive mechanism through a small angular increment. Probe motion along the baseline is then repeated and another step in baseline orientation is taken. These motions continue until the baseline has rotated 180°. The final result is the sample grid in the Fourier plane shown in Figure 7-6. The half-wave spacing indicated ensures that the Nyquist criterion is met.

The process of sampling along a baseline is often called "aperture-synthesis." The terms "rotation synthesis" or "supersynthesis" are reserved for the sampling procedure using the rotating baseline. In the literature these operations are treated merely as successive portions of the Fourier-plane imaging process.
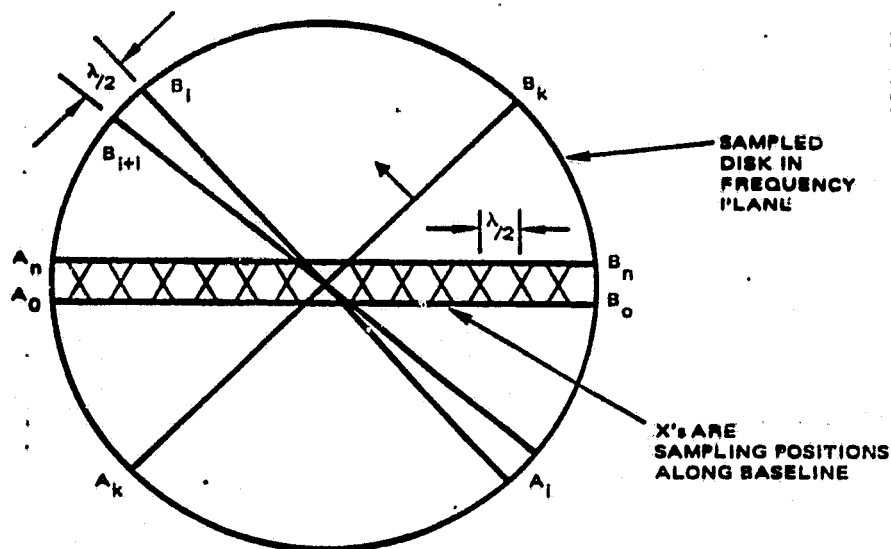
FIGURE 7-6. STRUCTURE OF FREQUENCY PLANE IMAGER SAMPLE GRID

8. BASIC APERTURE SYNTHESIS SYSTEMS

8.1 RESPONSE OF AN INTERFEROMETER TO AN EXTENDED, NONCOHERENT SOURCE

This section is not intended to be a mathematical analysis in depth of aperture-synthesis, since this has been done many times before (Reference 70, 71). The intent is to make a more heuristic examination backed by appropriate references. Comment will also be made about the applicability of the technique to the problem of orbital radiometric mapping.

The response of the aperture-synthesis interferometer to an extended source is analyzed in Appendix B. There it is shown that only an incoherent source produces a linear interferometer response. Eq. (B-16) of Appendix B shows this response, $P(X)$, to the power distribution over the scene, $P_o(X)$, is

$$P(X) = \overline{\beta} P_o(X) * K(X) \qquad (I\text{-}8\text{-}1)$$

where $K(X)$ is the antenna power pattern of the interferometer. Then Eq. (B-20) shows the relation between the space-domain and frequency domain functions first introduced in Section 6, Antennas as Spatial-Frequency Filters. Namely,

$$
\begin{array}{ccccc}
P(X) & = & P_o(X) & * & K(X) \\
\Updownarrow & & \Updownarrow & & \Updownarrow \\
S(F) & = & O(F) & \cdot & \overline{MTF}(F)
\end{array}
\qquad (I\text{-}8\text{-}2)
$$

where $S(F)$ and $O(F)$ are the complex spatial frequency spectrums of the image and object.

Let $G_1(X)$ and $G_2(X)$ be the antenna field patterns of the two interferometer antennas. Let $A_1(\eta)$ and $A_2(\eta)$ be the excitation functions of the probe

apertures. Then Eq. (B-21) of Appendix B shows the relation of K(X) and MTF(F) to these four functions as

$$K(X) = G_1(X) \cdot G_2^*(X)$$
$$\big\Updownarrow \qquad \big\Updownarrow \qquad \big\Updownarrow$$
$$MTF(F) = A_1(\eta) \star A_2^*(\eta)$$

$$(I-8-3)$$

As explained earlier, Kraus uses the correlation theorem to replace the convolution by a correlation to obtain

$$\tilde{K}(X) = \tilde{G}_1(X) \cdot G_2^*(X)$$
$$\big\Updownarrow \qquad \qquad \big\Updownarrow$$
$$MTF(F) = A_1(\eta) \star A_2^*(\eta)$$

$$(I-8-4)$$

Graphic construction of the MTF per Eq. (I-8-4) was shown in Figure 6-7, and is repeated here for easy reference.

Now let us vary S_λ in Figure 8-1 from zero to a maximum value D. From Eq. (I-8-2) it may be seen that the interferometer acts as spectrum analyzer as it scans over the spatial frequency spectrum of the scene O(F). The output is the image spectrum, S(F).

Often the interferometer is treated as an inferior antenna which is ambiguous due to response of its many lobes to the object scene. Methods are then devised to resolve the ambiguities. The effort to resolve the ambiguities is made because the interferometer is much lighter and less costly than a filled aperture whose main beam equals the interferometer lobe width. Figure 8-2 shows the variable baseline interferometer as a spatial-frequency spectrometer.
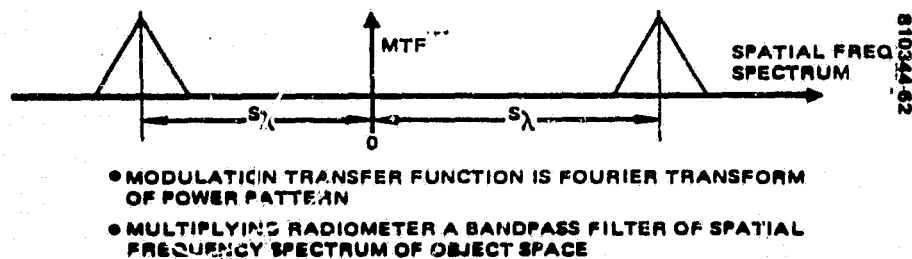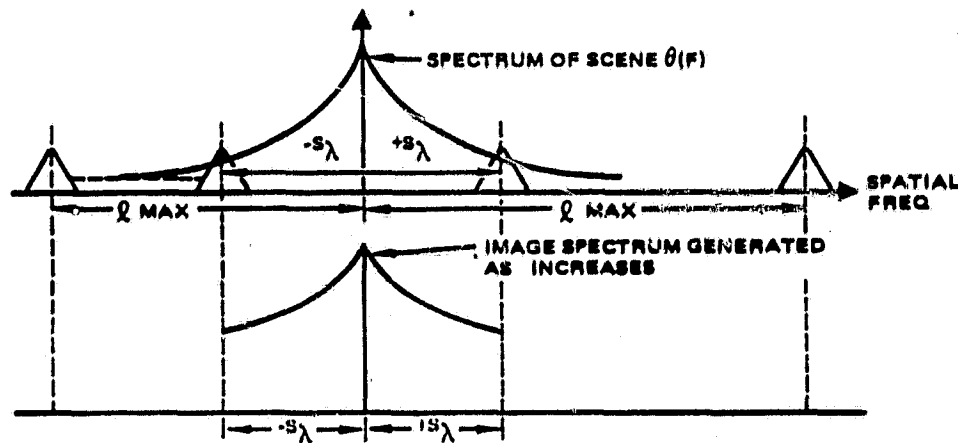

- MODULATION TRANSFER FUNCTION IS FOURIER TRANSFORM OF POWER PATTERN
- MULTIPLYING RADIOMETER A BANDPASS FILTER OF SPATIAL FREQUENCY SPECTRUM OF OBJECT SPACE

FIGURE 8-1. INTERFEROMETER MTF

SPECTRUM OF SCENE $\theta(F)$

$-S_\lambda$ $+S_\lambda$

SPATIAL FREQ

ℓ MAX ℓ MAX

IMAGE SPECTRUM GENERATED AS INCREASES

$-S_\lambda$ $'S_\lambda$

● INTERFORMETER IS SPECTRUM DISSECTOR OR SPATIAL-FREQUENCY SPECTRUM ANALYSER

● IN CONTRAST, FILLED APERTURE IS IMAGE DISSECTOR; ANALYZES SCENE 1 PIXEL AT A TIME BY ANGLE SCANNING

FIGURE 8-2. THE VARIABLE BASELINE INTERFEROMETER AS A SPATIAL-FREQUENCY SPECTROMETER

It is obvious from previous discussion that, when operated in the frequency domain, the interferometer is not ambiguous at all. In fact it is the pencil-beam antenna which is ambiguous.

Let the filled-aperture width be D, so that the pencil-beam and interferometer lobe have the same width. Then the low-pass MTF of the filled-aperture has a passband which extends over the entire spectrum out to frequency D. All the components of the spectrum are passed resulting in complete ambiguity.

The conclusion is that both the interferometer and the pencil-beam scanning antenna are both useful, unambiguous mappers when operated in their proper domains.

8.2 DISCUSSION OF LATERAL DECORRELATION EFFECTS

The above analysis of interferometer operation as a spatial-frequency spectrum analyzer may be confusing to those who have read radio astronomy textbooks and papers where the response of the interferometer to small uniform sources is analyzed. It is stated that when the source gets too big there is little response. A source subtending an angle of $1/S_\lambda$ is too big.

Kraus in his Figure 6-15 on page 175 shows why this is so by plotting the response of an adding interferometer to sources with three different angular extents. When the source becomes equal to or larger than the lobe spacing, the average of the source times the portion of the lobe pattern

falling on the source is nearly constant. There is then no usable output signal. The behavior for a multiplying radiometer is similar, except that DC term is not present. Kraus' Figure 6-14 is also given, since it shows how the point source response formed by multiplying the interferometer angular response by the antenna pattern of the interferometer antennas.

From Figure 8-3 it may be seen that as the baseline length increases and the lobes become thinner and more closely packed, the response of the interferometer to a finite source will decrease. For a uniform circular source of radius ρ located on the perpendicular bisector of the baseline at distance R, the normalized interferometer response is

$$\mu_{12}(S_\lambda) = \frac{2J_1(\nu)}{\nu} e^{i\psi} \qquad (I\text{-}8\text{-}5)$$

where ψ is the phase difference between radiation at the ends of the baseline and

$$\nu = 2\pi S_\lambda \frac{\rho}{R} = \pi S_\lambda \Delta\alpha \qquad (I\text{-}8\text{-}6)$$

The angle subtended R by the source is $\Delta\alpha$.

(a) INDIVIDUAL-ELEMENT PATTERN; (b) ARRAY PATTERN; AND (c) THE RESULTANT INTERFEROMETER PATTERN FOR THE CASE OF POINT SOURCE.

INTERFEROMETER PATTERN (a) FOR POINT SOURCE; (b) FOR A UNIFORM EXTENDED SOURCE OF ANGLE $\alpha < 1/s\lambda$; AND (c) FOR A UNIFORM EXTENDED SOURCE OF ANGLE $\alpha = 1/s\lambda$
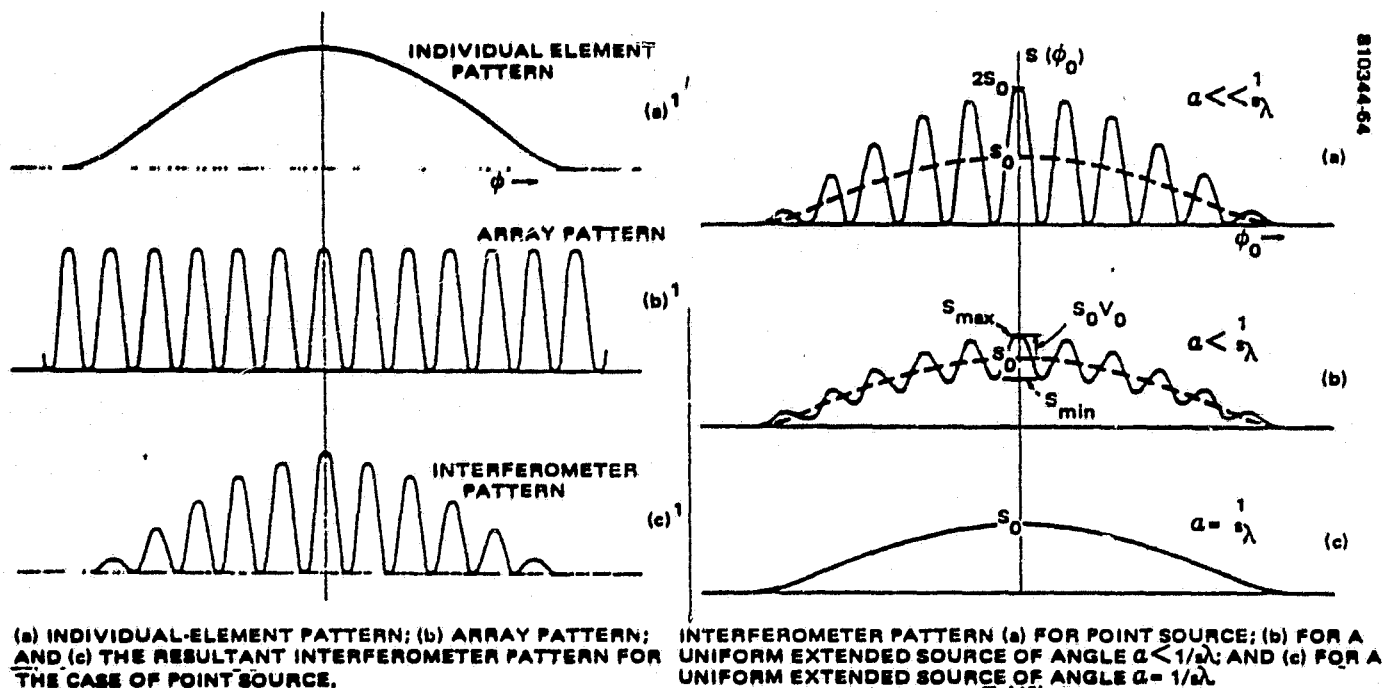
FIGURE 8-3. RESPONSE OF AN INTERFEROMETER TO SOURCES OF VARIOUS SIZES

In Figure 8-4 the plot of $|\mu_{12}|$, due to Born and Wolf (Section 10.4.3, pg 512, Figure 10.7) is shown.

The abscissa is not ν, but stems from an illustrative optical experiment. The first null occurs for $\nu = 3.83$. Let us assume that the interferometer output is decorrelated when $|\mu_{12}|$ falls to 3 db below its peak value. Since $|\mu_{12}|$ is the amplitude of the correlator output, -3 db occurs when $|\mu_{12}| = 1/\sqrt{2}$. At this point $\nu = 1.6$.

Substituting in Eq. (I-8-6) and solving for $\Delta\alpha$, we get

$$\Delta\alpha = \frac{1}{2S\lambda} = \frac{\Delta\theta}{2} \qquad\qquad (I-8-7)$$

Eq. (I-8-7) says that the disk cannot have an angular extent greater than half a lobe width without suffering significant decorrelation.

It may appear hard to reconcile these results with the earlier explanation of the interferometer performance in terms of a narrow bandpass filter which scans the spatial frequency spectrum of the scene. However these calculations concerning the decorrelation of a disk source are in truth a help in understanding the earlier analysis.
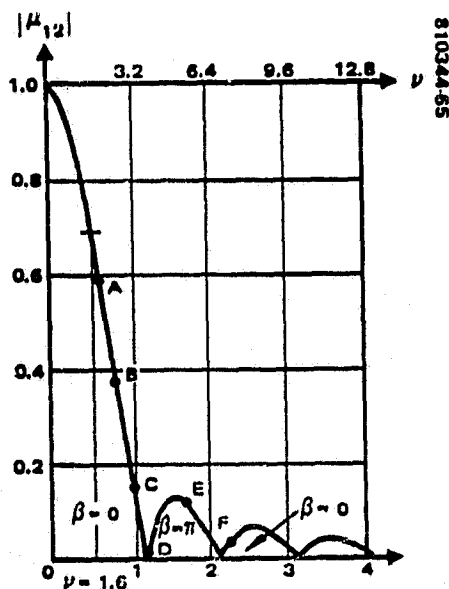
FIGURE 8-4. NORMALIZED LATERAL COHERENCE FUNCTION OF A NONCOHERENT CIRCULAR DISK SOURCE

The key to this puzzle is the fact that the terrain is not a featureless surface like the disk. It is composed of structures of all sizes ranging from continental dimensions down to facets only a wavelength or so in size. The facets can be vegetation, wave facets, ice patterns, soil and rock patterns and many other things, including emissivity variations over the surface of man-made objects. Let us model this terrain structure as a surface over which a multitude of disks are strewn. The disks have random emissivities clustered around the average terrain radiant brightness.

Let the disk diameters have a uniform probability distribution ranging from the operating wavelength up to the diameter of the mapped area. Disks larger or smaller than these limits cannot be seen. Call these limit diameters d_{min} and d_{max}. Also let the interferometer be located at distance R from the surface. The angles subtended by the disks range from the angular field-of-view downward to λ/R. That is

$$\alpha_{max} = \overline{FOV}, \qquad \Delta\alpha_{max} \approx 1 \text{ rad}$$

$$\text{(I-8-8)}$$

$$\alpha_{min} = \lambda/R, \qquad \Delta\alpha_{min} \ll 1 \text{ rad}$$

Aperture synthesis starts with the travelling probe very near the reference probe. In order for the largest disks to contribute to the output of the interferometer correlator, Eq. (I-8-7) requires that

$$S_{\lambda_{min}} \leq \frac{1}{2\Delta\alpha_{max}} \qquad \text{(I-8-9)}$$

when $S_\lambda = S_{\lambda_{min}}$ all the disks contribute. As S_λ increases the contributions of the larger disks drop out.

Let $X = S_\lambda/S_{\lambda_{min}}$; then $1 < X < \infty$. Then the disk size corresponding to X is

$$\alpha(x) = \left(\frac{1}{2S_{\lambda_{min}}}\right)\left(\frac{S_{\lambda_{min}}}{S_\lambda}\right) = \frac{b}{x} \qquad \text{(I-8-10)}$$

Assume that all disks smaller than $\alpha(X)$ correlate completely and all disks larger than $\alpha(X)$ do not correlate at all. The contribution to the correlation function of a disk will be proportional to $\alpha^2(X)$. Therefore

$$\vartheta(X) = \int_X^\infty \alpha^2(y)dy = b^2 \int_X^\infty \frac{dy}{y^2} = -b^2 \frac{1}{y}\Big|_X^\infty = \frac{b^2}{X}, \quad 1 \leq X \leq \infty \qquad (I-8-11)$$

We will set $\vartheta(X)$ equal to zero in the region it cannot be measured, namely zero to one. Then from Vol. I, Sec. 1.2, Eq. (3) of the Bateman Tables we find that the cosine Fourier transform of (8-11) is $[Ci(y)]$. Therefore

$$\left.\begin{array}{l} f(X) = 0, \quad 0<X<1 \\ \\ f(X) = \frac{1}{X}, \quad 1<X<\infty \end{array}\right\} \quad \xrightleftharpoons{} \quad -Ci(y) \qquad (I-8-12)$$

Figure 8-5 shows a plot of $Ci(y)$ taken from the Bureau of Standards Tables (pg 232, Figure 5.6). This is the spectrum associated with the correlation function (lateral coherence function) that was just derived.

The spectrum does not extend below $y = 1$ because $\vartheta(X)$ was not measured for X less than ± 1. This portion of the correlation function carried the information about the low frequency portion of the spectrum. This gap in $\vartheta(X)$ also produces the oscillatory behavior of the spectrum plot. The oscillations would not be present in a more realistic model. Such a model would be more random than the one used here.
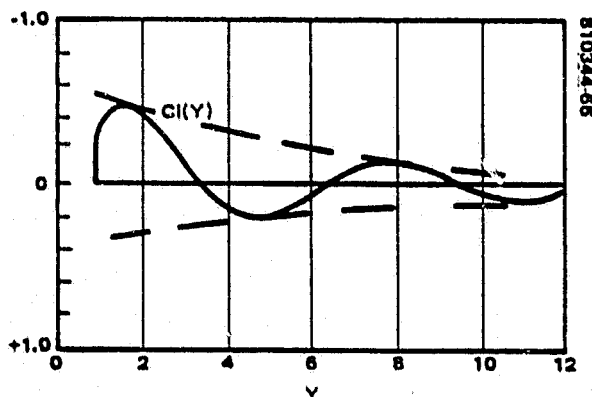

FIGURE 8-5. PLOT OF THE COSINE
INTEGRAL FUNCTION

However the absolute values of the maximums and minimums of Ci(y) act as envelope which shows the behavior of the more random model. From pg 244 of the Bureau of Standards Tables we find that

$$\left|Ci(y_{peak})\right| = \left|Ci[(n + 1/2)\pi]\right| \approx \frac{1}{y_{peak}} - \frac{2}{y_{peak}^3} \qquad (I-8-13)$$

Remembering that $\vartheta(X)$ is represented as a voltage at the output of the correlator, it may be seen that the spatial frequency spectrum is of the form

$$S(F) \approx F^2 \{\vartheta(X)\} \approx \left|Ci(X_{peak})\right| \approx \frac{1}{y_{peak}^2} \approx F^{-2} \qquad (I-8-14)$$

This result is not too far different from the spectrum shape selected in Appendix III on the basis of other considerations. This spectrum was $S(F) = F^{-2.5}$.

8.3 DETERMINATION OF MAXIMUM BASELINE LENGTH

In most of the literature the maximum baseline length is treated as a specified parameter. A perfect mapping system is considered to be one having a MTF which is unity out to the critical frequency of the system. Bracewell defines the signal emerging from such a system as the principal solution. Since there is lots of room in orbit, a less fatalistic attitude will be taken. It will be assumed that the baseline can be as long as we care to make it. As the baseline lengthens the spatial frequency spectrum of the scene will be measured with higher and higher resolution. This is illustrated by Figure 8-6 due to Steel. (Reference 72)

As spectrum resolution increases the rms difference between the signal representing the true scene and the signal produced by the mapping radiometer decreases. This rms error is one which does not decrease as the integration time available increases. In statistics an error of this type is called a bias error. The bias squared, b^2, is the integral of this video power density of the scene from D to infinity.

However as the maximum baseline length increases, the number of spectrum samples increases. Since this maximum length is the radius of the disk over which samples must be taken, the required number of samples increases as the square of the maximum baseline length, D. The total observation time available is fixed by the mapping geometry and the satellite velocity. Therefore the time available increases as D^2 as the maximum baseline length increases.
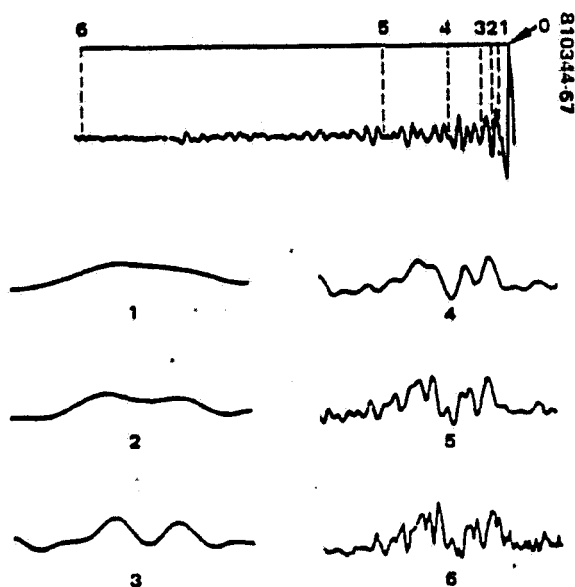
FIGURE 8-6. PROGRESSIVE INCREASE IN
SPECTRUM RESOLUTION WITH LENGTHENING
OF THE CORRELATION RECORD

The total mean square error is the sum of ΔT^2 and b^2 added power-wise. It is obvious that this sum goes through a minimum for some value of D. Call this optimum baseline length D_{opt}. This is a common type of statistical calculation. Figure 8-7 shows such an error plot for a temporal filter due to Papoulis. (Reference 73)

Oppenheim and Shafer (Reference 74) show plots of a spectrum where the total signal has been chopped up into records with lengths of 14, 51, 137 and 452 samples. The total number of samples is 3,164. So there are about 226, 62, 23 and 7 records, respectively, having the above lengths. A spectrum
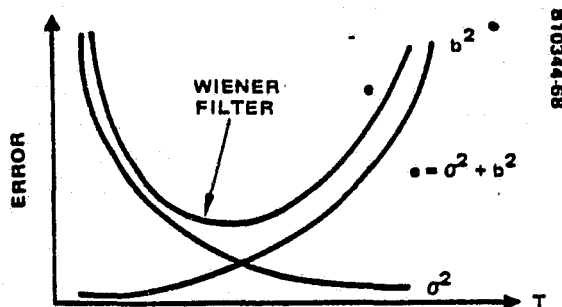

FIGURE 8-7. PLOT OF NOISE VARIANCE, ΔT^2,
BIAS SQUARED, b^2, AND TOTAL MEAN SQUARE
ERROR, c^2 FOR A TEMPORAL FILTER

estimate is made for each record; then all the records of a given length are averaged together. The results are shown in Figure 8-8 (Oppenheim and Shafer, Figure 11.3).

Here the noise is the result of the fact that the variance of a spectrum estimate obtained by Fourier transforming a correlation function approaches a non-zero constant as $N \to \infty$. The limiting variance of a spectrum estimate such as this was given in Eq. (I-5-8) of this report.

It may be seen that N = 51 is near to being the optimum record length.

In Appendix G, "Performance of Aperture-Synthesis Imagers," the optimum disk diameter as well as error expressions are derived. In Appendix XI, "Recapitulation of Radiometric Mapping Performance," plots of the rms noise error (square root of the variance) and the bias error are given for two aperture-synthesis systems. One system is in geosynchronous orbit. The other system is in low earth orbit. Plausible parameters are used in both cases. We have for GEO: $\Delta T = 30^\circ K$ = rms terrain fluctuation, $D_{opt} = 1.8$ km, $f = 1.2$ GHz, ST = $3^\circ K$ = temp res, b = $6.25^\circ K$ = bias, $\epsilon = 7^\circ K$ = total error. We see that for LEO: $D_{opt} = 1.5M$, f = 20 GHz, $\delta T = 3^\circ K$, b = $8^\circ K$, c = $9^\circ K$. These temperature resolutions are poor compared to mapping radiometers being designed at the present time. This occurs because many designs ignore bias error completely, both bias due to imperfect spectrum reproduction of spatial frequencies less than D and the finite-resolution error which is the present subject of discussion.
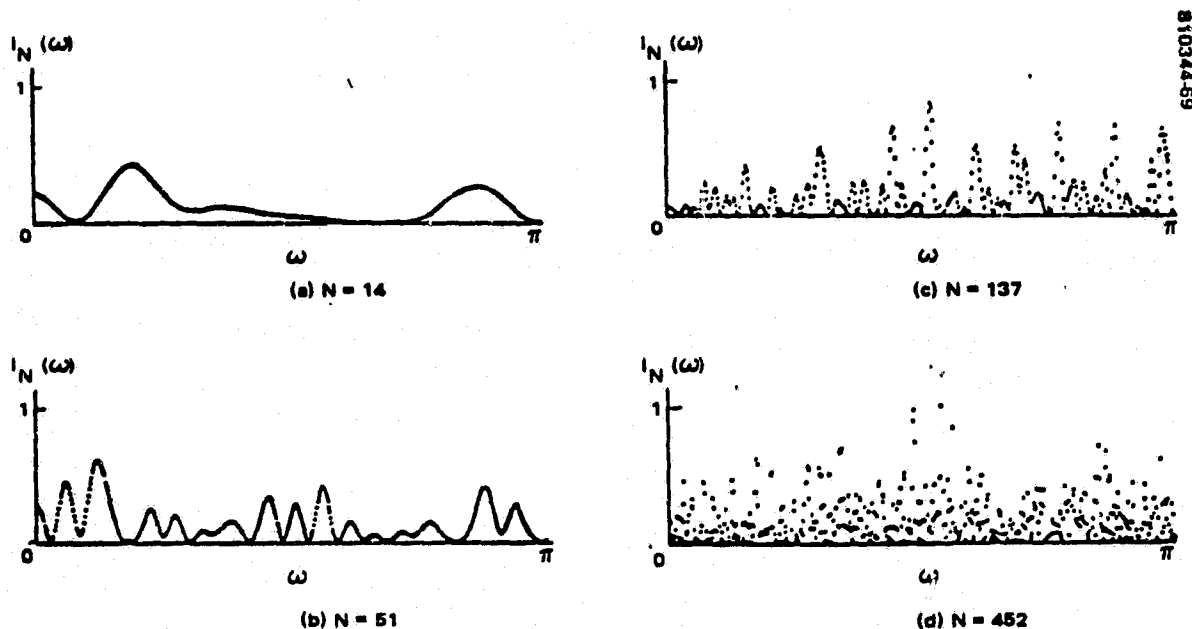

FIGURE 8-8. PERIODOGRAMS FOR A FIXED TOTAL NUMBER OF SAMPLES AND RECORDS OF LENGTH N

Probably the designer knows intuitively that he cannot build an antenna large enough to reach optimum size, although this does not excuse failure to consider bias errors in the frequency range of 0 to D. This portion of the spectrum is the signal. Distortion of this portion of the spectrum is distortion of the signal.

In many cases the system is used in applications which do not require resolution of detail smaller than some critical size. Then bias is found by integration from frequency D out to a frequency corresponding to this size, rather than out to infinity. This will decrease D_{opt} and also the corresponding bias and noise errors.

A second comment about this situation is that once bias error is acknowledged the system can be designed to reduce the error. Aperture synthesis does not turn in too good a performance record because bandwidth is limited by decorrelation effects. Figures 9 and 10 of Appendix show that true-time-delay antennas do better. But the optimum 12M antenna in LEO is certainly at the limits of the state of the art; the optimum 3KM antenna in GEO is certainly impossible. However SPINRAD (Figures 5 and 6) and RADSAR (Figures 7 and 8) can be built with optimum size. They also turn in fair radiometric performance ($\delta T \approx 1.5 - 2^{O}K$ and $b \approx 2.5 - 3^{O}K$). Contemporary design trends in radiometry such as Hadamard transform techniques (Reference 75) will probably reduce the errors values just quoted considerably in the future.

8.4 BIAS DUE TO FINITE RESOLUTION OF THE SPATIAL FREQUENCY SPECTRUM

Up to now we have only been concerned with bias due to finite angular resolution. There are other bias errors which are present and should be mentioned. One bias is due to the finite bandwidth of the bandpass MTF of the interferometer. In Figure 8-2 it may be seen that as the MTF scans over S(F) the result is the convolution of the two functions. The multiplicative relation, $S(F) = \vartheta(F) \cdot MTF(F)$ holds for a fixed baseline length. Let $\hat{S}(F)$ be the estimate of S(F) generated by the scanning MTF of the interferometer as the baseline length is changed. Then

$$\hat{S}(F) = \vartheta(F) * \overline{MTF}(S_{\lambda}) \qquad (I-8-15)$$

This result is completely analagous to the situation encountered when a temporal spectrum is measured by a scanning temporal filter. The bias for this case is found by Bendat and Piersol (Reference 76). It is

$$b[\hat{S}(F)] = \frac{B_e^2}{24} S''(F) \qquad (I-8-16)$$

where B_e is the filter bandwidth.

I-8-11

The Fourier transform of the MTF passband shape is the element pattern of the two probe antennas alone. The interferometer pattern appears only when the pattern origin is made the center of the baseline and the probe patterns have phases referred to that origin. The power pattern for the interferometer is then generated as shown in Figure 8-4. The element pattern is given by

$$E(X) = G(X) \cdot G^*(X) \tag{I-8-17}$$

The relations (8-2) for a fixed baseline are therefore replaced by the relations in (8-18). $\hat{P}(X)$ is the estimate of the intensity distribution function for the scene.

$$\hat{P}(X) = P_o(X) \cdot E(X)$$

$$\hat{S}(F) = \vartheta(F) * MTF(S_\lambda) \tag{I-8-18}$$

Note that $\hat{P}(X)$ is quite a good reproduction of $P_o(X)$, if the probe antennas have an aperture size which is small compared to the maximum baseline length, D. Then $\overline{MTF}(S_\lambda)$ is nearly an impulse. Also $E(X)$ is nearly uniform across the FOV subtended by the scene, except near the edges.

It would appear that $P_o(X)$ could be recovered by dividing $\hat{P}(X)$ by $E(X)$. However this inverse filtering operation will amplify noise greatly near the edges of the scene where $E(X)$ is small. The solution to this difficulty is use of a Wiener filter, rather than an inverse filter. The Wiener filter weights the degree of inversion carried out by the signal-to-noise ratio. This will be discussed in more detail in a later section.

It might be mentioned that the response of the aperture synthesis system over the spatial frequency band from zero to D is almost rectangular. This is much closer to perfect rectangular response than the response of a filled aperture of size D (shown in Figure 6-1). As a result the aperture synthesis system has almost twice the resolution of a filled aperture having the same size.

8.5 ROTATION SYNTHESIS

As mentioned in Section VII, "Frequency-Domain Imagers," the lateral coherence function is measured over a disk in the Fourier plane by rotating the interferometer baseline while repeating the aperture synthesis procedure. These repetitions are made often enough to meet the Nyquist criterion. The

Van Cittert-Zernike theory says that the two-dimensional Fourier transform of the lateral coherence function is the source brightness distribution (Appendix E).

There is a second relation between Γ_{12} and B which is less obvious. Let Γ_{12} ($\vec{\tau}$) be the lateral coherence function measured over displacement τ. Then the one-dimensional transform of Γ_{12} ($\vec{\tau}$) is the brightness distribution of the scene measured along a cut through $S(\upsilon, \nu)$ which has the same orientation as the displacement $\vec{\tau}$ in the frequency plane. This is the projection theorem.

This theorem shows that B(x, y) can be built up out of a series of 1-d transforms of Γ_{12} measured along a set of diameters of the Fourier disk.

No matter how B is computed from Γ_{12}, the process of measuring Γ_{12} along a set of disk diameters in the frequency plane followed by computation of B(x, y) is called rotation synthesis.

8.6   ALTERNATE WAYS OF SCANNING THE INTERFEROMETER MTF IN SPATIAL FREQUENCY

Up to now we have examined aperture synthesis where the length of the interferometer baseline has been varied. There are two other means of varying the effective baseline length.

The first alternate method is variation of the operating frequency. This changes the variable $S_\lambda = S/\lambda$ by varying λ instead of S. Sheridan (Reference 78) built a system of this sort for mapping the sun (Reference 78). It is apparent that the source must have the same emissivity over a very wide range of frequencies when this technique is used. Also, the radiometer must be located where it will not receive radiation from man-made source. This is a very hard condition to meet.

The MFASR avoids this problem by using a discrete set of frequencies which fall at places in the radio spectrum where there is little interference. The question then arises as to whether there are enough samples available to permit a reasonable reconstruction of the image. This problem is treated in later sections.

The other possible method of varying S_λ is rotation of the interferometer end-over-end. This is illustrated in Figure 8-9.

The axis of rotation is at right angles to the axis of rotation for rotation synthesis. If the field of view is centered on the zenith-nadir line ZN, the effective interferometer baseline length equals D cos θ. This varies from a value of D when the baseline is horizontal to zero when it falls on ZN.

The field-of-view must be held fixed by rotating the radiators on the baseline ends in the opposite direction from the base rotation. This is indicated in the drawing by the circular arrows marked -θ.
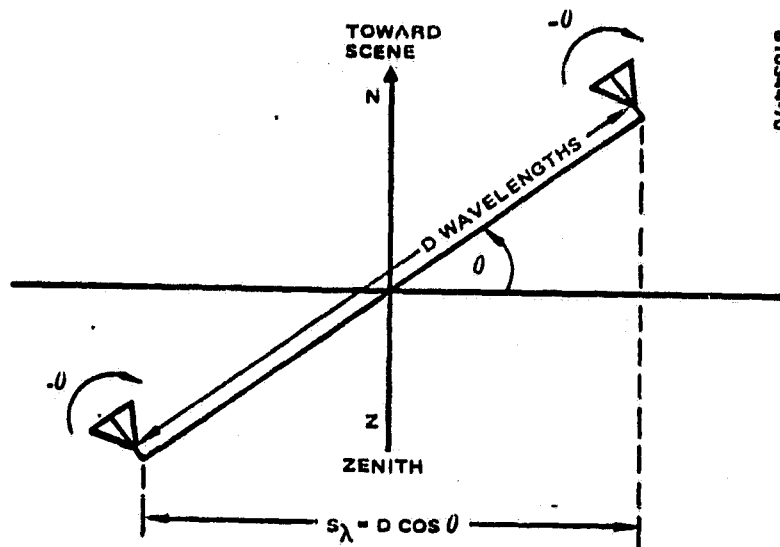
FIGURE 8-9. ROTATING INTERFEROMETER BASELINE

9. IMAGE FORMATTING AND OPTIMIZATION

9.1 INTRODUCTION

Frequency domain imagers tend to invite a much larger amount of data processing than conventional imaging systems. In the case of conventional systems, the image is formed by hardware. Image-data processing is then an additional step. If the system designer elects not to take this step, the system still delivers images to the user.

In the case of frequency-domain images, data processing is mandatory. Even if nothing else is done, the spectrum data must be Fourier transformed in order to produce the image. Once data is digitized and fed to a computer in order to effect the frequency-to-space-domain transition there is strong motivation for additional data processing in order to improve image quality.

Image manipulation by computer tends to quantify image quality to a degree which never occurs when a lens or pencil-beam scanner forms the image. This quantification of image properties leads to definition of the operations on the image data which are needed to produce the best possible images. This in turn leads to specification of the performance of the rest of the imaging system required to get the best results out of the image-data processing.

Although there are grey areas, image processing can be broken down into three areas. One area I will call formatting. This involves arrangement of the image data for best display. Examples are image annotation, dynamic range adjustment to match the capabilities of the display, contour line generation, and map-projection generation.

A second area is image manipulation in order to improve quality according to some criterion.

A third area is measurement of image parameters and comparison between images in order to quantify image quality.

Fortunately, there is a software section VICAR which meets most routine image processing needs. The rest of this chapter and Volume II are concerned with image processing associated with specific Fourier-plane imager systems.

9.2 THE SOFTWARE SYSTEM - VICAR

This software is described by Castleman (Reference 117) in his Chapter 4 and Appendix K, VICAR PROGRAM INDEX. Appendix M includes a library of FORTRAN language image-processing programs.

· ·The VICAR system, and its documentation, may be obtained from the following.

> COSMIC
> Computer Center
> 112 Barrow Hall
> University of Georgia
> Athens, GA 30601
> (404) 542-3265

The VICAR system is filed under number NPO-13415. A set of image-processing programs for the PDP-11 computer are also filed there under NPO-14892 and the name Mini-VICAR.

Appendix K is a reproduction of Castleman's Appendix II where a list of VICAR programs are given together with brief descriptions. Asterisks indicate that the program was available from COSMIC at the time Castleman's book was published.

9.3 IMAGE QUALITY CRITERIONS

The most obvious criterion for image quality is the root-mean-square difference between the final image and the original scene. The total rms difference consists of two components. One component is system noise. Its magnitude is given by ΔT, the system temperature resolution. The second component is called bias.

In statistics bias is defined as error which does not go to zero as the averaging time increases without limit. One source of bias, the finite resolution of the system, has already been discussed in Section 8.4. There a method of minimizing the sum of the system noise variance and the square of the finite resolution bias was presented. As the linear array length or maximum interferometer baseline length increases system noise increases. This occurs because the number of spectrum samples required increases. However, the resolution bias decreases as the system resolution improves. It was found that there was an optimum array length where the sum of the noise variance and bias squared went through a minimum value.

Radiometer mapping performance for this optimum configuration was computed for three types of mapping systems in appendices to this report. These three mapping types are aperture synthesis, pencil-beam scanners using series fed antennas and true-time-delay antennas.

There are additional bias errors due to causes other than finite resolution. They will be discussed later in this chapter. A common type of bias error is nonuniform response of the system MTF between zero and the cutoff spatial frequency. That is the system does not produce the principal solution. In Bracewell's terminology the principal solution is obtained when the mapping system exhibits a rectangular MTF from zero out to the cutoff frequency.

Bias error due to motion of the scene is an important error type. It will be found that such motion produces a nonrectangular MTF which must be multiplied into the chain of transfer functions. This idea will be developed in more detail in a little while. Corrections for motion effects is therefore carried out as part of the process where deviation of the system MTF from rectangular is corrected.

The optimum filter for carrying out this correction is the Wiener filter. It will be discussed in the next section. If the image is to be used for visual inspection, rather than measurement, it will be found that the Wiener filter emphasizes low frequencies too much. The Wiener restored image appears dull and flat, rather than crisp. Some of the finer detail appears blurred.

In Section 5.8, the idea was introduced that the optimum image for visual inspection maximizes the area under the system MTF. The MTF of the eye is included in the chain of MTF's whose product is the system MTF. A filter which produces this result is the parametric Wiener filter (geometric mean filter). This filter, as well as other filters for improved image appearance, are discussed by Hunt (Reference 118) and by Stockham (Reference 119). Both papers give examples of the improvement of image quality compared to Wiener-filtered images.

A final criterion of image quality is the entropy of the image. Information theory shows that the entropy of a signal record is a measure of the information it contains. Therefore, if the conversion of the image signal from its spectrum form to the final space-domain image can be carried out so as to maximize image entropy, it can be argued that the best possible image results.

Presumably, this "best" image can now be processed so as to meet any other quality criterion since it contains more data about the original scene than any other possible image.

The importance of these observations lies in the fact that algorithms exist which can maximize image entropy in the process of reconstructing the image from spectral data. These algorithms go under the appellation, "maximum entropy method," or MEM. They are very important when the image must be reconstructed from undersampled spectrum data.

Since this is the problem we face in mechanizing the MFASMR system, these algorithms are treated in detail in Volume II of this report. Image reconstructions using MEM are presented there. Only a brief introduction to these methods will be given in this chapter.

The last type of optimum filter to be considered is the matched filter. This filter exhibits the maximum possible response when driven by the input waveform it was designed to match. The filter response does not resemble the input waveform in any way. Instead, the output is a pulse which approaches a delta function as the time-bandwidth product of the waveform it matches approaches infinity. This filter maximizes the probability of detecting the matching waveform in the presence of interfering noise.

9.4 WIENER FILTERS

If the modulation transfer function of the radiometer system is not rectangular, it would appear that all that had to be done to make it rectangular would be to add a filter to the MTF chain whose MTF is the reciprocal of the MTF of the rest of the system. This filter is called an inverse filter.

Unfortunately, in regions where the system MTF is zero the inverse filter transfer function becomes infinite. Figure 9-1 shows that even when the system MTF has no internal zero regions, noise becomes dominant at the high frequencies where the system is tending toward zero. The result is a noisy image.
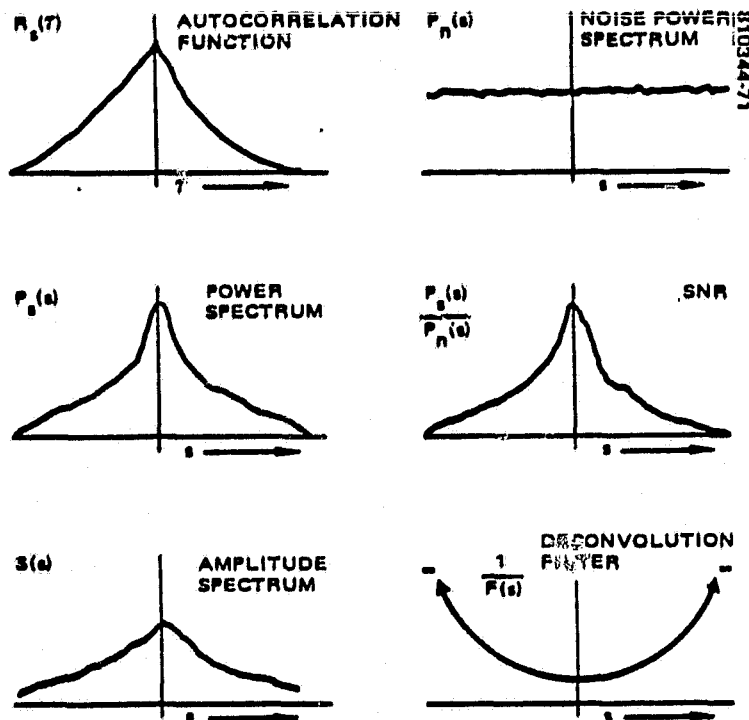

FIGURE 9-1. NOISE PROBLEM IN DECONVOLUTION

A correlation filter can be designed where the signal is large and little inversion where noise dominates. The transfer function of this filter is

$$G_W(u,v) = \frac{F^*(u,v)}{|F(u,v)|^2 + P_n(u,v)/P_S(u,v)}$$

(I-9-1)

This is the Wiener filter. It minimizes the rms difference between the original scene and the final image. Figure 9-2 shows the functions involved in Eq. (I-9-1) for the one-dimensional case.

A closely related filter is the power spectrum equalization filter. This filter restores the power spectrum of the degraded image to what it was before degradation. Its transfer function is

$$G_{PSE} = \left[\frac{P_S(u,v)}{|F(u,v)|^2 P_S(u,v) + P_N(u,v)} \right]^{1/2}$$
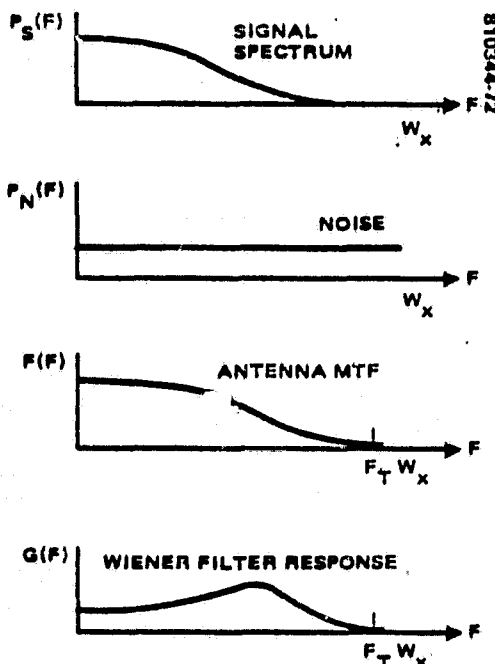
(I-9-2)

FIGURE 9-2. DETERMINATION OF THE
WIENER FILTER RESPONSE

Like the Wiener filter, the PSE filter is phaseless (real and even). Both filters reduce to straight deconvolution in the absence of noise, and both cutoff completely in the absence of signal.

However, the PSE filter (sometimes called a homomorphic filter) does not cutoff at internal zeroes of the point-spread function $F(u,v)$.

The other filter related to the Wiener filter is the parametric Wiener filter. It is of the form

$$G_{PW}(u,v) = \left[\frac{F^*(u,v)}{|F(u,v)|^2}\right]^{\alpha} \left[\frac{F^*(u,v)}{|F(u,v)|^2 P_S(u,v) + P_N(u,v)}\right]^{(1-\alpha)} \qquad \text{(I-9-3)}$$

This is a generalization of the filters already discussed. When $\alpha = 1$ Eq. (I-9-3) becomes the inverse filter. If $\alpha = 0$ and $\gamma = 1$ we have the Wiener filter. When γ is not unity the result is a parametric Wiener filter. The constant γ may be selected for any desired amount of Wiener-type smoothing. When $\gamma = 0$ we are back to the straight deconvolution filter. If $\alpha = 1/2$ and $\gamma = 1$, Eq. (I-9-3) defines a filter which is the geometric mean between ordinary deconvolution and Wiener deconcolution. The parametric Wiener filter with $\gamma < 1$ and the geometric mean filter with the same constraint define the range for γ where the best visual images are obtained.

In the construction of error budgets needed to judge the suitability of a system for a given application we need to be able to compute the residual rms error after Wiener filtering. Both signal and noise power density spectrums enter into the determination of this error. The error expression is

$$\epsilon^2 = \int_0^\infty \int_0^\infty \frac{P_S(u,v)P_N(u,v)}{P_S(u,v) + P_N(u,v)} \, du\,dv \qquad \text{(I-9-4)}$$

This is a generalization of two dimensions of an expression given by Bendat (Reference 120) for the rms error of a Wiener filter.

Now Eq. (3), Appendix G gives the signal spectrum in polar coordinates as

$$P_S(\rho,\theta) = \frac{\delta T^2}{4\pi\sqrt{FOV} \, \rho^{5/2}} \qquad \text{(I-9-5)}$$

where δT is the rms radiance fluctuation of the object scene, $\overline{FOV}$ = field-of-view ≤ 1 rad.

For a uniformly illuminated antenna the normalized transfer function is $(1-\rho)$, $0 \leq \rho \leq 1$. Let the noise spectrum be flat with typical rms amplitude of 0.5 to $1^\circ K$. Since δT varies from $20^\circ K$ for rural areas to $40^\circ K$ or so for urban areas, ice/water or land/water images, $(\Delta T/\delta T)^2$ is much less than one. So in polar coordinates we have

$$\epsilon^2 = \int_0^1 \int_0^{2\pi} \frac{P_S(\rho,\theta)(1-\rho)\Delta T^2 \rho}{P_S(\rho,0)(1-\rho) + \Delta T^2} = 2\pi\Delta t^2 \int_0^{1-\Delta\rho} \frac{\rho d\rho}{1 + \left[4\pi\sqrt{FOV}\rho^{5/2}/(1-\rho)\left[(\Delta T/\delta T)^2\right]\right]}$$

(I-9-6)

As long as ρ is a little less than 1 the second term of the denominator is small. Therefore,

$$\epsilon^2 = 2\pi\Delta T^2 \frac{\rho^2}{2} \Big|_0^{1-\Delta\rho} = \pi\Delta T^2$$

(I-9-7)

Therefore, the total rms error is approximately $\Delta T\sqrt{\pi}$.

It is important to note that Wiener filtering cannot be used directly for image reconstruction when the spectrum is undersampled. This difficulty arises because missing data samples are represented by zeroes. The Wiener filter will cut off in any spectral region with zero spectral power density. An estimate of the original spectrum with undersampling eliminated must be generated by interpolating in the regions where samples are missing. Some sort of two-dimensional spline function is needed.

If polynominal splines are used, the polynominal coefficients must be determined so that the spline gives known values of the spectrum where samples exist. As an alternate we can use an expansion into a set of orthonormal functions. As in the polynominal spline case, the series coefficients are picked to make the expansion give the correct values at points where samples exist.

If Wiener filtering is carried out after the spectrum has been filled in by interpolation, it is not obvious that we still have a system which minimizes rms error. The performance of Wiener filtering when the input data is undersampled is treated in more detail in Volume II where computer results are reported.

9.5 MAXIMUM ENTROPY RECONSTRUCTION METHOD

It will be found in the next chapter that the amount of distortion calculated for the MFASMR is probably too great to make the system very useful. These calculations are based on a discipline called "rate distortion theory." This theory assumes that the image reconstruction process is linear. However, there are more general non-linear methods which are more powerful than linear image construction.

As mentioned in the introduction to this chapter, one of the most prominent processes of this type is the maximum entropy method, MEM. This process is recursive. Rather than requiring a priori statistical data about signal and noise, MEM effectively collects the required statistical data during the recursion cycles. It goes without saying that MEM is nonlinear.

In order to determine whether MEM or similar techniques can reduce distortion in MFASFMR to a tolerable level, computer experimentation was used. This work is reported in detail in Volume II. Only a few comments will be made about MEM at this point in order to paint a complete picture of the image reconstruction problems - albeit a sketchy one.

An excellent place to start a survey of MEM techniques is the IEEE Reprint Volume edited by Childers (Reference 121). The promise of MEM is well illustrated by a comparison of spectrum estimates derived from the spectrum autocorrelation function made by Ables (Reference 122). Ables' plots in Figure 1 to 4 are reproduced here as Figure 9-3. It may be seen that the MEM spectrum estimate is considerably better than either the weighted or the unweighted Fourier transform of the autocorrelation function of the signal.
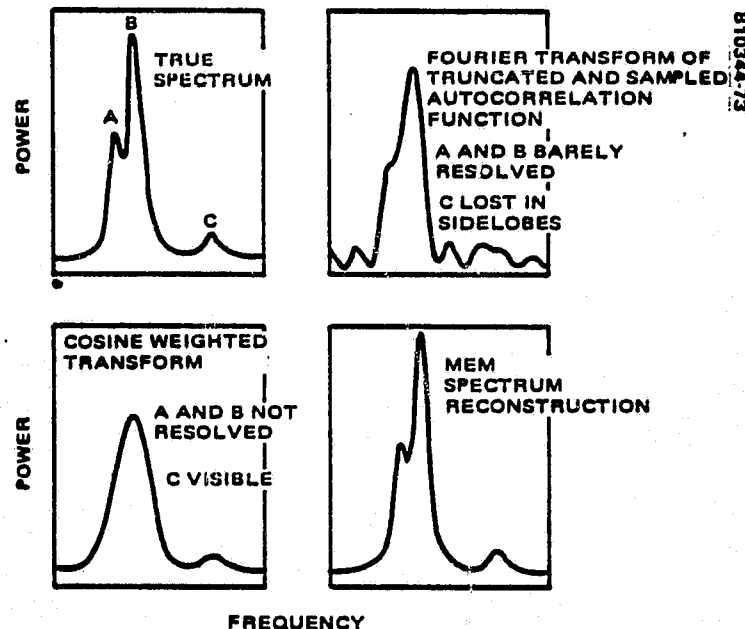

FIGURE 9-3. COMPARISON OF FOURIER – TRANSFORM
SPECTRUM ESTIMATE WITH MAXIMUM ENTROPY
SPECTRUM ESTIMATE

In spite of fine results obtained with MEM, the method has generated
considerable controversy. Some MEM algorithms have been unstable. As the
number of recursions increase some MEM algorithms produce artifacts. Typically,
these artifacts are sharp spikes which grow on the edges of spectrum lines as
recursion continues. An example of this behavior, taken from a report by
King (Reference 123), is shown in Figure 9-4. The fourth recursion is the best
estimate of the original spectrum. The spikes, which are well developed by
the seventh recursion, obviously increase the entropy of the signal. Apparently,
the MEM estimate increases entropy beyond that of the true signal by producing
artifacts.

However, Wernecke (References 124, 125, 126, 127) has developed a stable
MEM algorithm for image reconstruction. Stability is ensured by imposing a
constraint on the calculation of the estimated image during the recursion pro-
cess. This constraint prevents any point in the image estimate, being formed
during a recursion cycle, from taking on a negative value. This condition is
of course consistent with the nature of image signals, since such signals are
always positive everywhere.

This condition may not appear to be a very severe constraint. However,
it is sufficient to ensure stability of the MEM reconstruction process. Almost
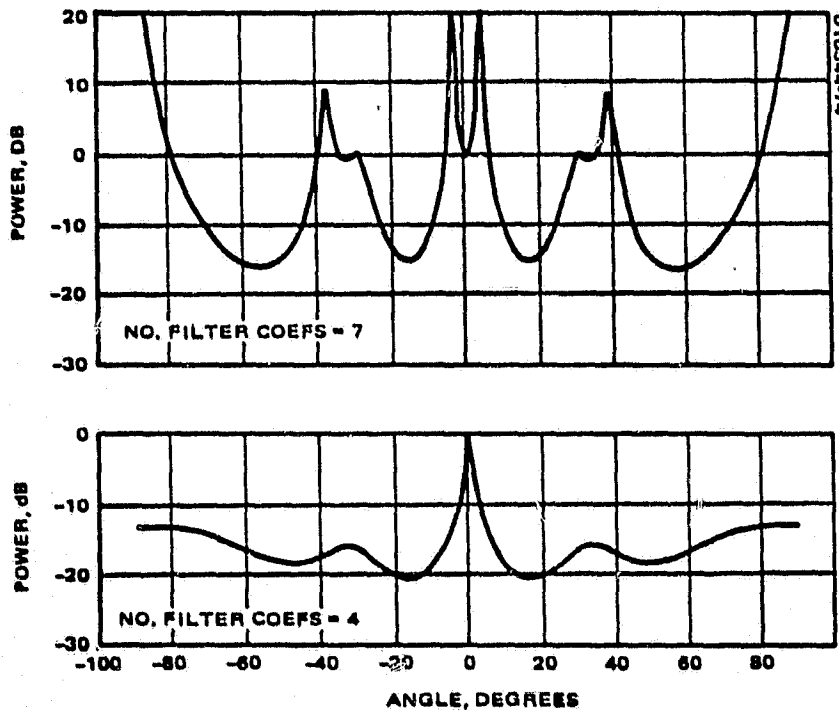

FIGURE 9-4. GROWTH OF ARTIFACTS DURING RECURSION BY
AN UNSTABLE MEM ALGORITHM

all possible unstable processes involve oscillatory behavior which grows with time. This oscillatory disturbance drives part of the image estimate negative. Constraining the image estimate to be always positive damps out this oscillatory behavior. Then recursion cycles of all orders are stable. As recursion continues the entropy of the images estimates asymptotically approaches a maximum value.

Volume II will shows results for real radiometer image-data. However, a "sneak preview" will be given here where the original object scene is two cubes stacked on top of each other. This is similar to a test object used by Wernecke in Reference 124. Wernecke's "stacked blocks" object is shown in his Figure 3.7. In Figure 9-5, which follows, the stacked blocks phantom used by Dr. Ming Chang is shown. Dr. Chang carried out the computer study of image reconstruction for MFASMR reported in Volume II. The following data comes from this study.

Following Figure 9-5, Figure 9-6 shows the amplitude spectrum of the stacked blocks object. Figure 9-7 shows the residual spectrum after samples are taken over the spectrum plane as shown in the upper right hand corner of the figure. We now have an undersampled spectrum.

Figure 9-8 shows the reconstructed image obtained by taking the two-dimensional inverse Fourier transform. As one can see, the resulting image is quite poor.

Figures 9-9, 9-10, and 9-11 show the MEM image reconstruction process taking place. The results of 2, 4, 6, 10, 18, 22, and 25 iterations are shown.
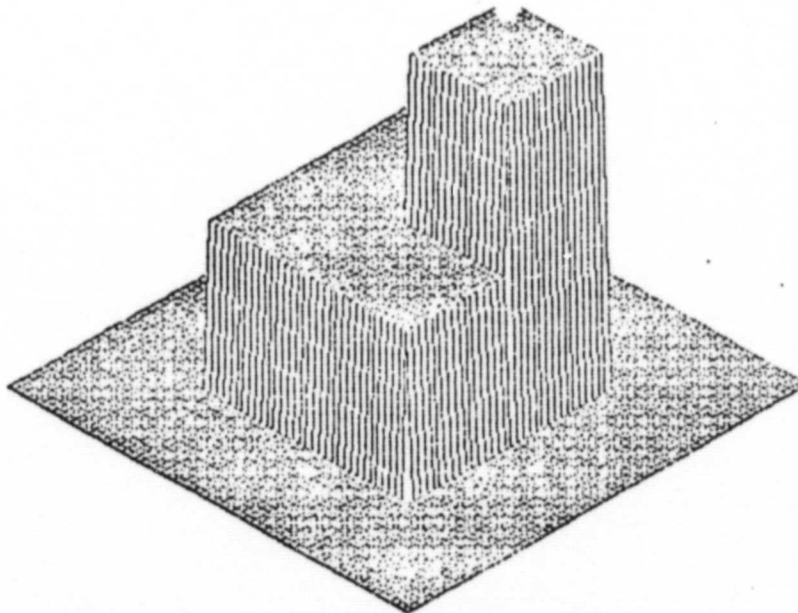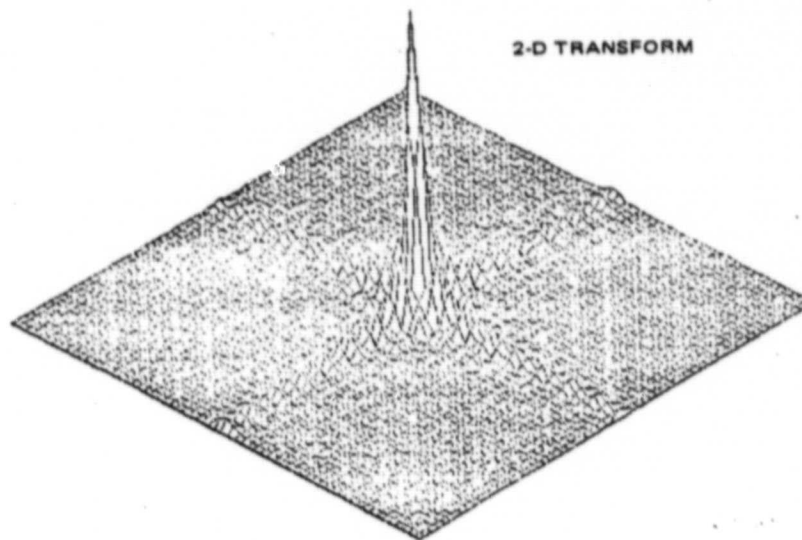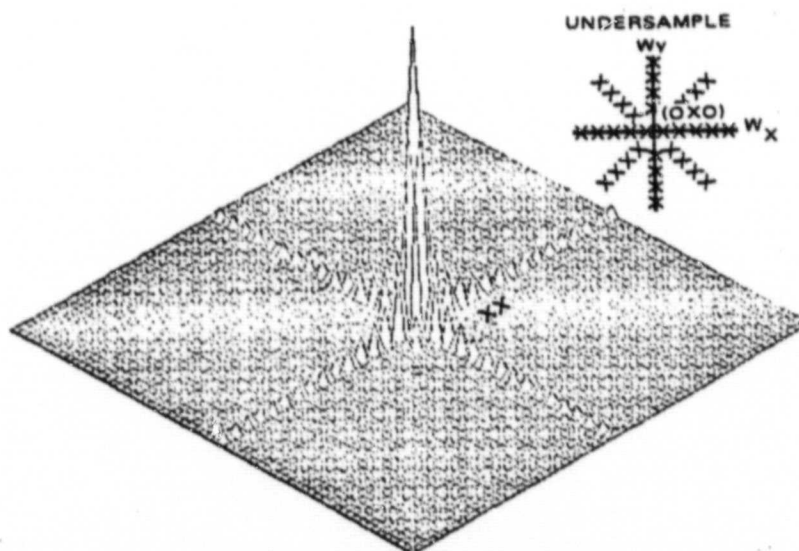

FIGURE 9-5. STACKED BLOCKS PHANTOM OBJECT

2-D TRANSFORM

810344-76

FIGURE 9-6. AMPLITUDE SPECTRUM OF STACKED
BLOCKS OBJECT

UNDERSAMPLE

810344-77

FIGURE 9-7. STACKED BLOCKS RESIDUAL OBJECT
SPECTRUM AFTER UNDERSAMPLING

DIRECT FFT-1
RECONSTRUCTION

810344-78

FIGURE 9-8. IMAGE OF STACKED BLOCKS OBTAINED BY
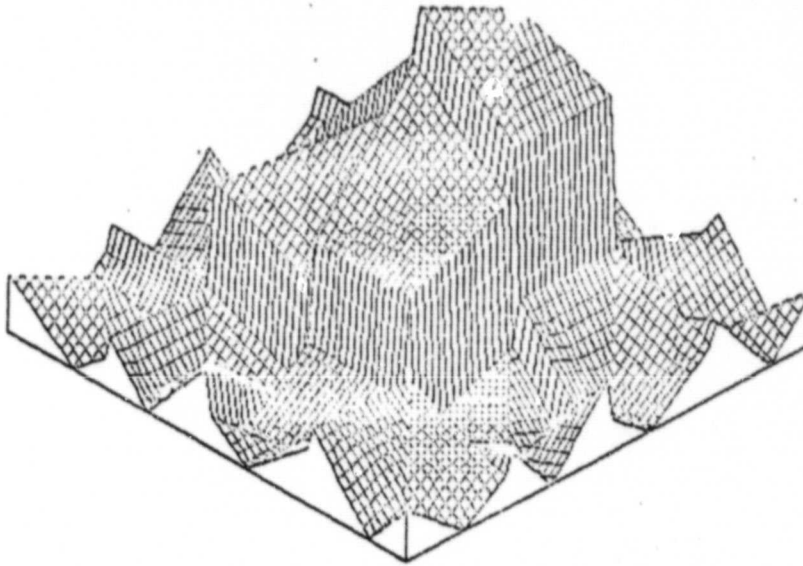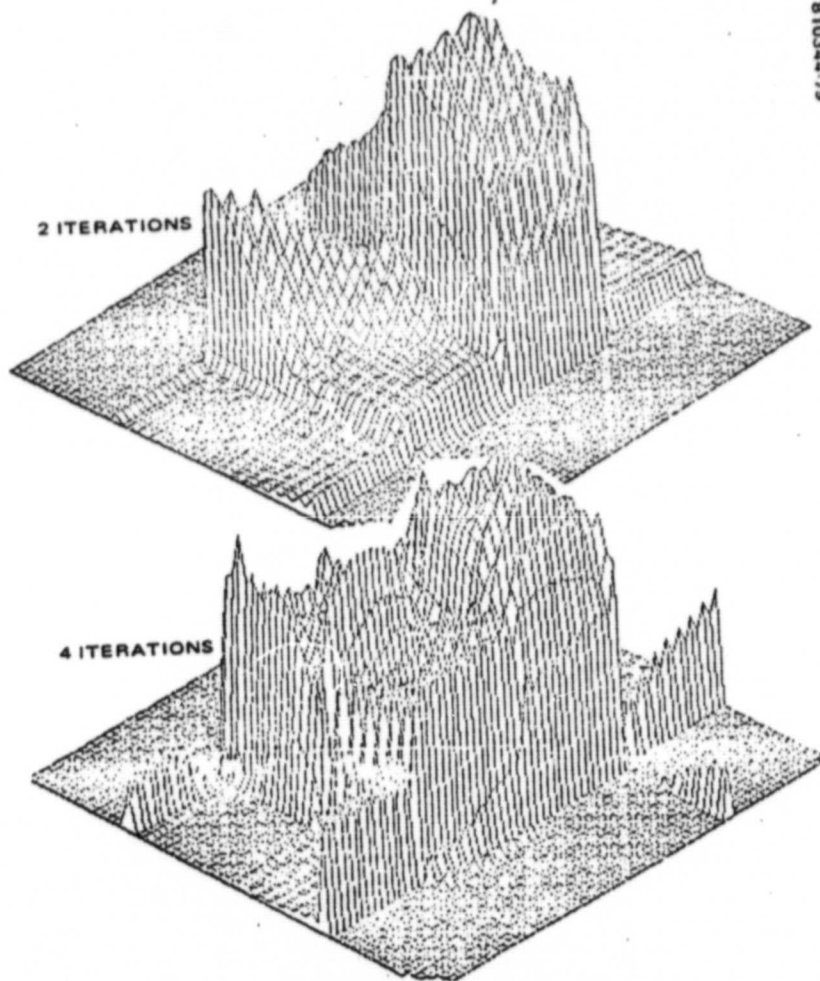INVERSE FOURIER TRANSFORM OF UNDERSAMPLED
SPECTRUM

810344-79

2 ITERATIONS

4 ITERATIONS

FIGURE 9-9. MEM RECONSTRUCTION

6 ITERATIONS

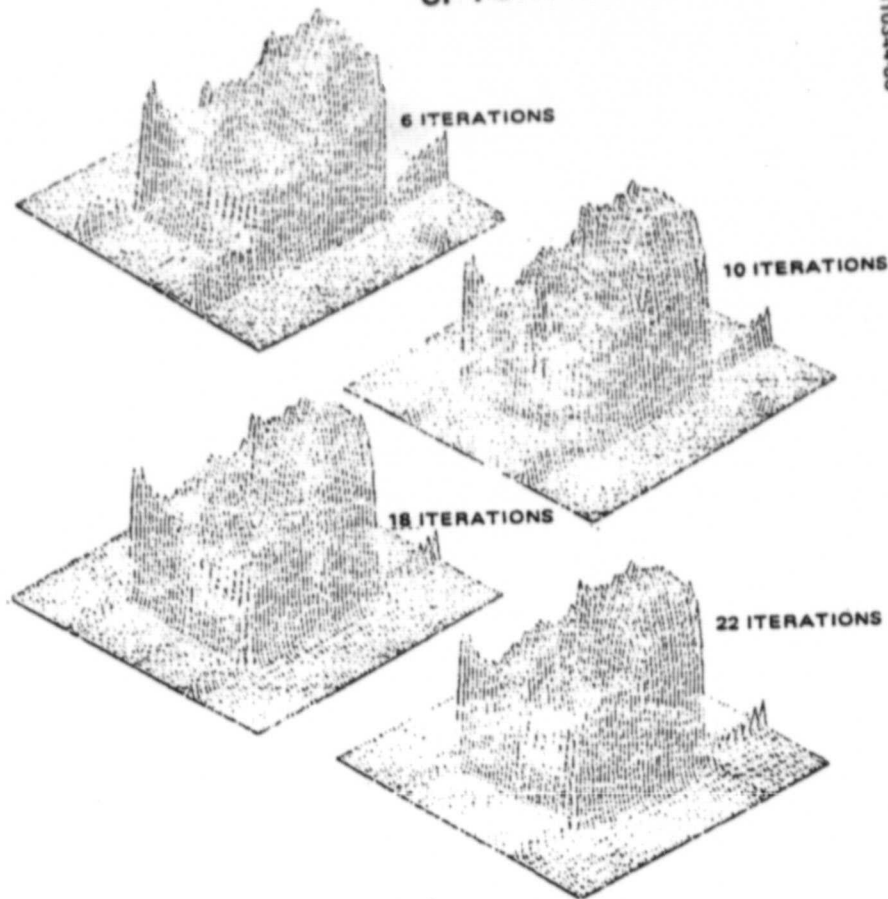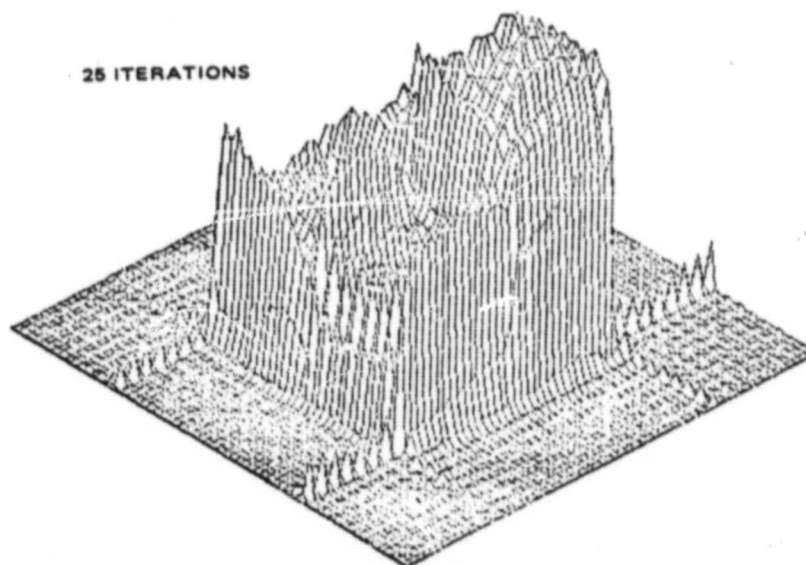10 ITERATIONS

18 ITERATIONS

22 ITERATIONS

FIGURE 9-10. MEM IMAGE RECONSTRUCTION

25 ITERATIONS

FIGURE 9-11. MEM IMAGE RECONSTRUCTION

Although the image is far from perfect, after six to ten iterations, it is
considerably better than the direct inverse Fourier transform. The thin ridges
extending out from the cube corners are similar to the impulses which appear
at function discontinuities when the Fourier transform is taken without under-
sampling. This behavior is called Gibbs' phenomenon. Since this Gibbs arti-
fact occurs in both reconstruction methods, it does not appear reasonable to
fault the MEM algorithm for this generating this imperfection.

Continued iteration up to 25 recursions shows that the reconstruction is
stable. It does not have spikes which grow indefinitely as was the case for
King's MEM algorithm (Figure 9-4).

9.6   IMPACT OF MEM RECONSTRUCTION METHODS ON SYSTEM DESIGN

Even when there is no undersampling MEM reconstruction is an attractive
method of forming the image. The MEM method is adaptive in that we do not
have to know a priori statistical data about signal and noise. The signal
statistics directly controls the MEM recursion as the image is built up.

As we have seen, Wiener filters require quite a bit of signal and noise
statistics. This is a severe drawback, since the mapped terrain hardly quali-
fies as a stationary source. The statistics vary as we look at farmland, for-
rest, the sea, ice fields, etc. There are extensions of Wiener filter which
can adapt the filter to a locally stationary signal (Reference 128). However,
these systems are complicated. MEM reconstruction appears to be the better
answer.

Now it stands to reason that if we are going to maximize the entropy of
the final image by use of MEM reconstruction, we should design the rest of the
system to produce a signal with maximum entropy rate at the input to the MEM
process. Appendix G, "Optimizing Antenna Illumination and Bandwidth for Maxi-
mum Radiometric Information Transfer," by Dr. Jack Gustinsic shows how an entire
radiometric mapping system can be designed to maximize signal entropy.

It was hoped that mapping systems, other than the series-fed array system
selected by Gustincic, could be treated in the same manner. However, there is
not time enough to do this at present. Also the size of this report, which is
probably too big already, would increase even more. However, it is possible to
comment on some general design principles.

First Gustincic found that maximum information rate occurred for approxi-
mately uniform antenna weighting. This was for a system using a series fed
array. In the other mapping systems considered, which use true-time delay
antennas or transform arrays, it turns out that exact uniform weighting is
optimum. Similarly, the spatial frequency spectrum measured by aperture syn-
thesis should be uniformly weighted in order to obtain maximum information rate.

The overall transfer function of the radiometer receiver, treated as a filter, is subject to a similar constraint. Goldman (Reference 129) shows that the entropy loss through a filter is zero only when the filter has a rectangular passband. The virtues of rectangular passbands for radiometers has been recognized for quite a while, even though improvement has been given as a decrease in temperature resolution rather than increase in information rate (Reference 130).

9.7 MOTION COMPENSATION

If the image data is not undersampled, there is one type of error which should be removed by Wiener filtering before MEM image reconstruction. This error is due to image motion. Removal of blur due to motion is very important when the mapper is in low-earth orbit.

Let us examine the signal at an image plane, such as a focal plane of a lens. This image can be decomposed into an assembly of points. Let these points move at a constant velocity v due to motion smear. Let the time required to sense the image be t. This can be the exposure time of a camera, the time required to measure the frequency spectrum of the scene or the time needed to complete a raster scan with a pencil beam. The displacement of the point during the observation time is $\Delta X = mvt$. The smeared point of brightness I is described by the expression

$$f(x) = \begin{cases} A, & -\frac{\Delta X}{2} \leq X \leq +\frac{\Delta X}{2} \\ \\ 0, & \text{otherwise} \end{cases} \tag{I-9-8}$$

where A is proportional to I. Thus each point is converted into a rectangle. The corresponding Fourier transform is

$$G_{MOTION} = A\Delta X \frac{\sin(F\Delta X/2)}{(F\Delta X/2)} \tag{I-9-9}$$

We can conclude from Eq. (I-9-9) that image motion introduces a sinc transfer function into the system chain of MTF's. It may be seen that, in conformance with earlier comments, the motion error cannot be removed by introduction of the inverse filter A/G(F). Its poles at $F = 2n\pi/\Delta x$ render it unstable. The "correct" image will be dominated by noise.

The removal of image blur by digital and electronic filtering has been treated by Sawchuk (Reference 131, 132, 133). Optical filters for removal of motion blur by optical processing is treated by YU (Reference 134) and by Tsujiuchi (Reference 135). In Figure 9-12 the rectangular response to the moving point is shown in a plot from Tsujiuchi's paper. The restoration of the point image is accomplished by use of a dual filter. One filter section
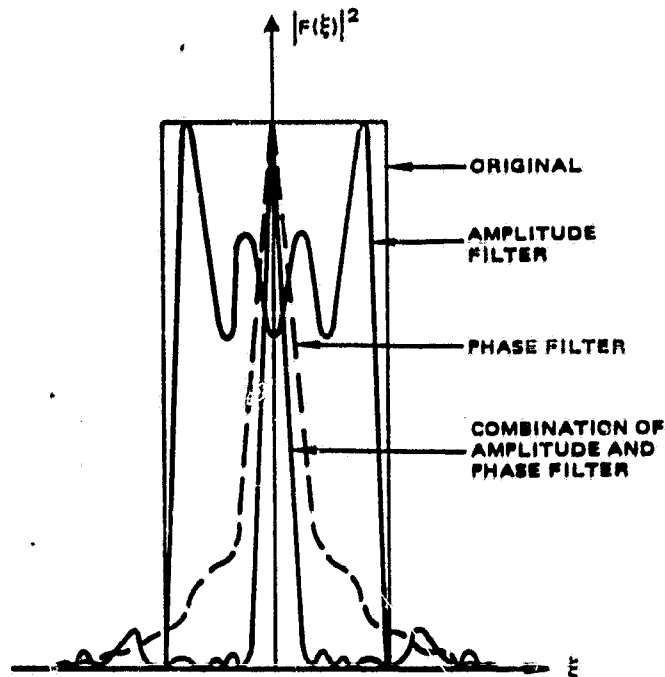
FIGURE 9-12. CALCULATED IRRADIANCE OF A
RESTORED POINT IMAGE BLURRED BY LINEAR
MOTION

modifies only amplitude; the other section modifies only phase. The plot shows the effect of each section separately and their combined effect. The smeared point is labeled original. The system impulse response to a point source is labeled "Combination Amplitude and Phase Filter."

In the following discussion it is important to note that the error was not produced in the sensor. It was the image data itself which is degraded. Let us examine the response of pencil-beam scanning imager to this degraded data.

Jacobs (Reference 136) computes the degradation of the spatial-frequency response of a scanning pencil-beam imager due·to image motion. He uses the product of the pencil-beam scan rate , and the radiometer time constant as a parameter which characterizes the rate of image motion. The scan rate is tied to image motion by the requirement that the scan process must be rapid enough to keep up with the image motion.

The degraded MTF due to scanning is found to be

$$MTF'(\omega) = \frac{MTF(\omega)}{\sqrt{1 + (\tau v w)^2}}$$

(I-9-10)

where MTF(ω) is the transfer function of the antenna and MTF'(ω) is the transfer function degraded by scanning. The degradation factor $[1 + (\tau v\omega)^2]^{-1/2}$ was plotted by Jacobs and shown below as Figure 9-13.

We find that pencil-beam scanners also suffer degradation due to image motion. For some reason, degradation due to image motion is usually not calculated when the system performance of such mappers is estimated. Also pencil-beam scanner systems almost never use restoration filters to compensate for motion. However, these systems are using the same motion-degraded images that frequency-plane imagers use. Degradation due to the motion is just as severe. Motion correction will be just as beneficial.

Consideration of these effects when evaluating both imager system types will lead to better comparisons of their performances.

9.8   BIAS DUE TO SPECTRUM CURVATURE

In Figure 8-2 the bandpass MTF of an interferometer is shown measuring the spatial frequency spectrum of the scene. In the next chapter Figure 9-4 shows how measurement of the spectrum of the signal received by a frequency-scanning array produces an image of the terrain.

In both cases these spectrums are measured by filters with finite passbands. A filter producing no error would give the power density of the spectrum at a frequency equal to the center frequency of the filter passband.
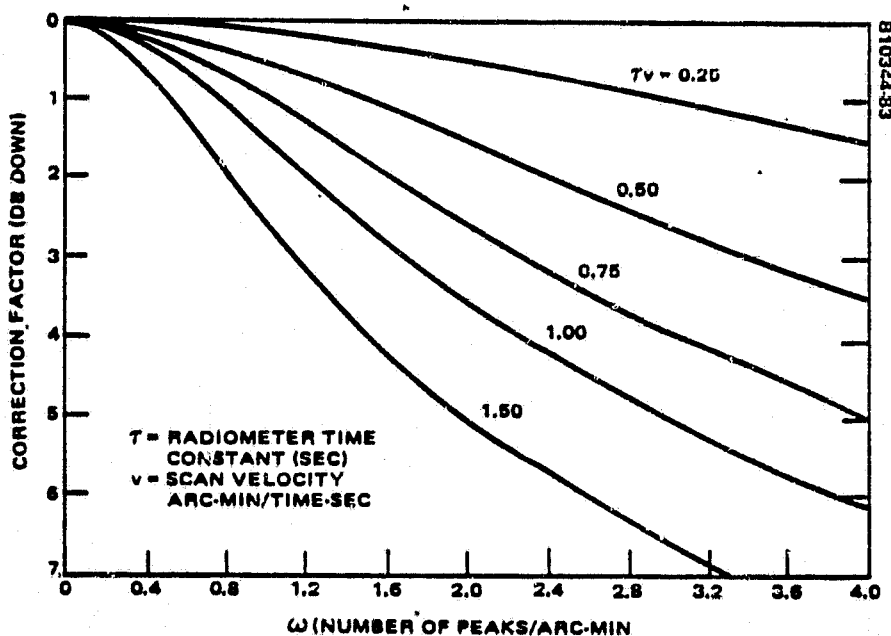

FIGURE 9-13. IMAGE DEGRADATION FACTOR AS A
FUNCTION OF NORMALIZED SCAN RATE

Due to spectrum curvature and the finite bandwidth of the spectrum measuring filters they do not find the midband spectrum density. The measurement is perturbed by a bias error. According to Bendat and Piersol (Reference 137) this bias error is given by

$$b^2[\hat{G}(f)] \approx \frac{B^2}{24} G''(f) \qquad\qquad (I\text{-}9\text{-}11)$$

where $G''(f)$ is the second derivative of the spectrum density function and B is the bandwidth of the measurement filter. There is no point in computing these biases for the raw spectrums. Their shape and curvature will be altered by filters used to correct for image errors in the image restoration process.

However, this bias should be computed along with the other system biases in the process of drawing up a bias budget for the purpose of estimating system performance.

10. UNDERSAMPLING AND IMAGE RECONSTRUCTION

10.1 INTRODUCTION

At this point, enough background material has been introduced to proceed with an analytic evaluation of the Multifrequency Aperture Synthesizer Microwave Radiometer (MFASMR). This frequency-imaging system varies the baseline length of its interferometers by stepping the operating frequency through a set of values scattered over most of the microwave spectrum. The frequencies selected are: 2.375, 4.75, 9.5, 19.0, and 38.0 GHz.

A typical MFASMR configuration is two interferometer baselines at right angles to each other. Baseline lengths of 10M with 1M dishes is also representative. Such a system is shown in Figure 10-1.

In Section 7.5, a paper by Ko, (Reference 68) "Coherence Theory in Radio-Astronomical Measurements," was referenced. In this paper, Ko shows that the maximum baseline length in wavelengths is the reciprocal of the system resolution in radians. Ko also states that, in order to prevent aliasing, the sample spacing in wavelengths is the reciprocal of the FOV in radians. Section 7.5 points out that this implies the moving interferometer antenna is stepped along the baseline in steps equal to the antenna width. When this observation
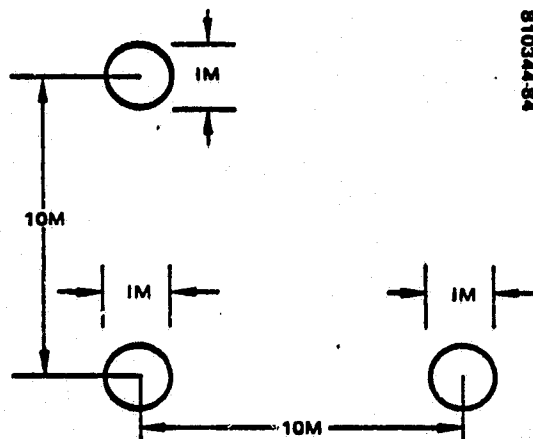

FIGURE 10-1. TYPICAL MFASMR CONFIGURATION

is applied to the configuration in Figure 10-1, we see that it generates a
10-line image (10 x 10 pixels), since the number of pixels in the final image
equals the number of complex spectrum samples.

Since five complex samples along the two baselines can be taken, using
five operating frequencies, we see that 10 samples out of a grid of 100 samples
can be taken. Only 10 percent of the raw image data is retained. The problems
associated with undersampling the spectrum of the scene dominate the processing
of the MFASMR data. As a result, it was decided to incorporate the analysis of
its performance in this chapter on undersampling and image reconstruction when
undersampling is present. Volume II describes a computer study of the same
problem. This study is more detailed than the analytic attack used here.

10.2   COMMENTS ON THE MFASMR ANTENNA CONFIGURATION

Before proceeding to MFASMR performance analysis, it is necessary to take
a look at the MFASMR antenna arrangement. This system is, of course, a fre-
quency-plane imager. The antenna arrangement, plus the designation of the
operating frequencies, determines the set of samples of the scene spectrum
available for image reconstruction.

The first comment concerns the use of two baselines at right angles to
each other. As mentioned in the Hughes proposal, this arrangement conveys
a strong feeling that the spatial frequency spectrum being sampled is a
separable function. A separable function is the product of two functions, each
of which is a function of a single independent variable. Let $F(U,V)$ be the
spectrum function. If it is separable, it is of the form $R(U)T(V)$, where $R(U)$
and $T(V)$ are known over two boundaries of the region in which $F(U,V)$ exists.
We know from theoretical physics that in order for separation to occur, the
set of eigenfunctions of R and T must have nodes over their corresponding
boundary surfaces. This rarely happens and then only for a restricted set of
coordinate systems. In such cases, specification of R and T over their
corresponding boundaries forces these functions to assume certain values
throughout the region occupied by the field, namely, values which are the sum
of the eigenfunction set with appropriate weighting.

It is easy to see that $F(U,V)$ fails this test. Specification of $F(U,V)$
along the interferometer baselines does not constrain $F(U,V)$ over the rest of
the frequency plane in any way.

A more formal proof that $F(U,V)$ is not separable stems from the projec-
tion theorem. This theorem showed that there is a one-to-one correspondence
between profiles generated by cuts through the space-domain function, say
$g(\theta,R)$, and corresponding frequency-domain profiles obtained by a cut through
$F(U,V)$, which has the same orientation as the cut through the object function,
$f(x,y)$. These two profiles are a one-dimensional Fourier-transform pair. The
above relation between $F(U,V)$ and $f(x,y)$ implies that we must measure $F(U,V)$

profiles over a set of orientations distributed over 180° with Nyquist sampling
density in order to be able to calculate f(x,y) exactly. These measurements
of the spatial frequency over a set of orientations is the process of rotation
synthesis, which by now has become familiar to us.

The minimum number of orientations which can be used with any hope of
image reconstruction are cuts in the U and V directions and along orientations
at ±45° with respect to the principal axes. This is the sort of sample set
used by Werneke (Reference 124) shown in Figure 3.8, page 96 of his doctoral
dissertation. The same sampling scheme has been used by Dr. Chang when
checking out Werneke's MEM program. It is shown in the upper right-hand cor-
ner of Figure 9-7 in this report.

In order to obtain such a sample set, another antenna must be added to
the original MFASMR configuration as shown in Figure 10-2.

The diagonal baselines AB and AD contribute absolutely essential additional
samples. Now the samples, although sparse, are indeed distributed over the
entire disk in the frequency plane corresponding to the scene. The potential
baselines DC and DB are not used. Since translation of a baseline without
orientation change does not affect the value of the mutual coherence function,
these baselines merely repeat the measurements made using baselines AC and AB.
We have added ten more samples to the ten available with the original con-
figuration. So we now have 20 samples out of a possible 100. The undersampled
spectrum is 20 percent of the total spectrum. This is also a statement that
the information rate of the undersampled data stream is 20 percent of the infor-
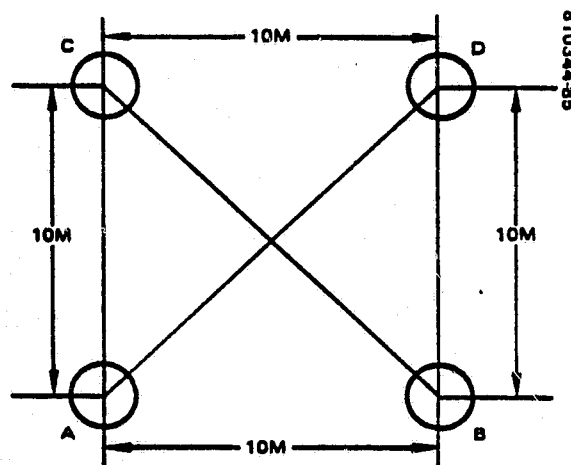mation rate of the sample stream associated with the Nyquist sampled spectrum.

FIGURE 10-2. MODIFIED MFASMR CONFIGURATION

10.3 RATE DISTORTION THEORY

Once the information rate of the coded image signal is known, we can ask what error is produced in the reconstructed image when the information rate is less than the information rate for Nyquist sampling. Theoretical techniques have been devised which find an upper limit on this error. These techniques go under the name of "rate distortion theory." (Reference 140)

The mean-square error introduced into the signal as a function of the rate of source coding and information transmission through a communications channel is best expressed in parameteric form. For this mean-square-error criterion and for a Gaussian source this parametric representation is analytically tractable. (References 141, 142)

The analysis assumes optimum coding. If the coding is optimum, then the distortion is an upper bound on the error for all sources with the same spectrum, but with other than Gaussian statistics.

The parametric form for the rate of source coding and for information transmission used the parameter θ. In Figure 10-3 it can be seen that θ is a decision level in a sorting process.

As indicated by the hatcher code, the portion of the power-density spectrum above the level θ is retained and transmitted. The crosshatched area, which can be interpreted as the frequency distribution of the error process, $x(t) - y(t)$, is often called the error spectral density. In ideal mean-square-error (MSE) encoding, no attempt is made to preserve any of the information contained in those regions of the frequency axis in which $\psi(\omega) \leq 0$, where $\phi(\omega) > \theta$, a fraction $[\phi(\omega) - \theta]/\phi(\omega)$ of the source power generated in an infinitesimally narrow band centered at ω is retained in the ideal reproduction. In other words, one should try to reproduce the portion of the source output generated at frequencies such that $\phi(\omega) > 0$ with a MSE of θ per unit bandwidth. For θ somewhat higher than illustrated in the figure (hence with distortion
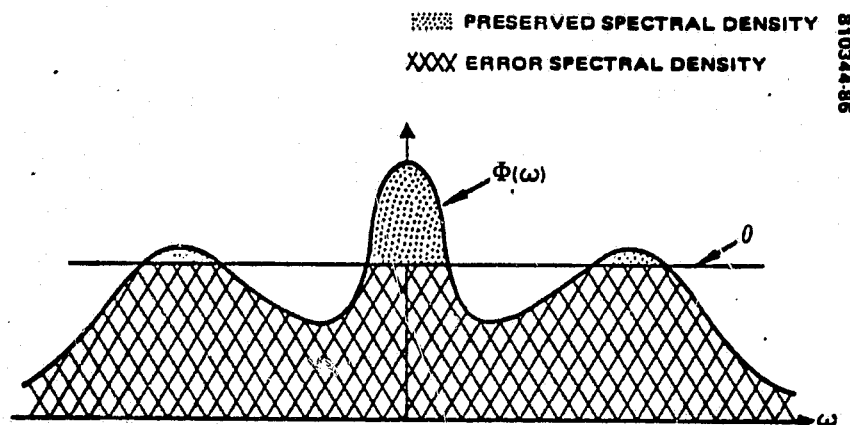

FIGURE 10-3. SPECTRAL DENSITIES OF SOURCE, ERROR, AND REPRODUCTION FOR OPTIMUM CODING

somewhat larger), the portion of $\phi(\omega)$ centered at $\omega = 0$ is the only one that an ideal system attempts to reproduce. For such θ one may put the process through an ideal low-pass filter prior to encoding. In general, a filter with several passbands and stopbands is required.

Then in terms of this parameter θ, from Berger, Eqs (4.5.51) and (4.5.52) the mean-square-error and information rate are

$$d(\theta) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \min \ [\theta, \phi(\omega)] d\omega \qquad (I-10-1)$$

$$R(\theta) = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \max \left[\theta, \ln \frac{\phi(\omega)}{\theta} \right] d\omega \qquad (I-10-2)$$

The variance of the process, which is used to normalize $d(\theta)$, is

$$\sigma^2 = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \phi(\omega) d\omega \qquad (I-10-3)$$

If the spectrum is one-sided, the integral limits of $\pm\infty$ are replaced by the limits 0 and $+\infty$.

In the two dimensional case the differential frequency element in the Fourier plane is of the form $(\rho \, d\phi \, d\rho)$. Starting with the Cartesian form of the two-dimensional integrals from Reference 142, Eqs (8.4.45) and (8.4.46), we see that in polar form all the integrals are of the form

$$I = \varsigma \int_0^{2\pi} \int_0^{\infty} f(\rho, \phi) \rho \, d\rho \, d\phi = 2\pi\varsigma \int_0^{\infty} f(\rho) \rho \, d\rho \qquad (I-10-4)$$

I-10-5

Before looking at our case let us examine some typical results from the literature. As an example, Berger uses a Gaussian source with spectral density of the form

$$(4.5.76) \quad \Phi(\omega) = A_n \left[1 + \left(\frac{\omega}{\omega_o} \right)^{2n} \right]^{-1}, \quad n = 1,2, \ldots \quad \text{(I-10-5)}$$

The equation number to the left is Berger's equation number.

Berger chooses A_n so that

$$\frac{1}{2\pi} \oint \Phi(\omega) \, d\omega = \sigma^2 \quad \text{(I-10-6)}$$

Letting $n \to \infty$ in Eq (I-10-5) yields an ideal Gaussian source with constant spectral density

$$(4.5.77) \quad \Phi(\omega) = \begin{cases} \pi\sigma^2/\omega_o, & |\omega| \leq \omega_o \\ \\ 0, & |\omega| > \omega_o \end{cases} \quad \text{(I-10-7)}$$

Substitution of this $\Phi(\omega)$ into Eqs (I-10-1) and (I-10-2) and elimination of θ yields the result

$$R(D) = \left(\frac{\omega_o}{2\pi} \right) \ln \left(\frac{\sigma^2}{D} \right) \quad \text{(I-10-8)}$$

where $R(D)$ is the information rate as a function of the MSE distortion, D. Hence, the MSE rate distortion function of a stationary white Gaussian source of average power σ^2 limited to frequency band $|\omega|/2\pi \leq B$ is given by Shannon's classic formula

$$(4.5.78) \quad R(D) = B \ln \left(\frac{\sigma^2}{D} \right) \quad \text{(I-10-9)}$$

Berger observes that the same R(D) curve applies to any stationary Gaussian process of average power σ^2 whose spectral density is flat over a set of positive frequencies of total measure B and vanishes elsewhere. In particular, it is important to note that the one-dimensional results also apply to a two-dimensional raster. Since there is no correlation from pixel to pixel for a white bandlimited Gaussian source, there is no correlation between raster lines. Therefore the raster lines can be unraveled to form a one-dimensional process of the type we have just been examining.

Berger plots the results of evaluating distortion as a function of rate and plots the results in his Figure 4.5.4 for the spectrum defined in Eq (I-10-5). This plot is reproduced below as Figure 10-4.

Note that as the normalized mean-square distortion approaches unity the information rate approaches zero. The white-noise source drops the most rapidly since it is the least correlated signal and therefore the most damaged by bandwidth compression.

Pratt (Reference 143) treats a common source model for image signals, the separable Markovian source whose elements possess equal variance σ^2 and possess adjacent pixel correlation factors ρ_R and ρ_C along rows and columns. He finds the rate distortion function for the one dimensional case to be

$$R_1(D) = \frac{1}{2} \log_2 \left[\frac{\sigma^2 (1 - \rho_R^2)}{D} \right] \qquad \text{(I-10-10)}$$
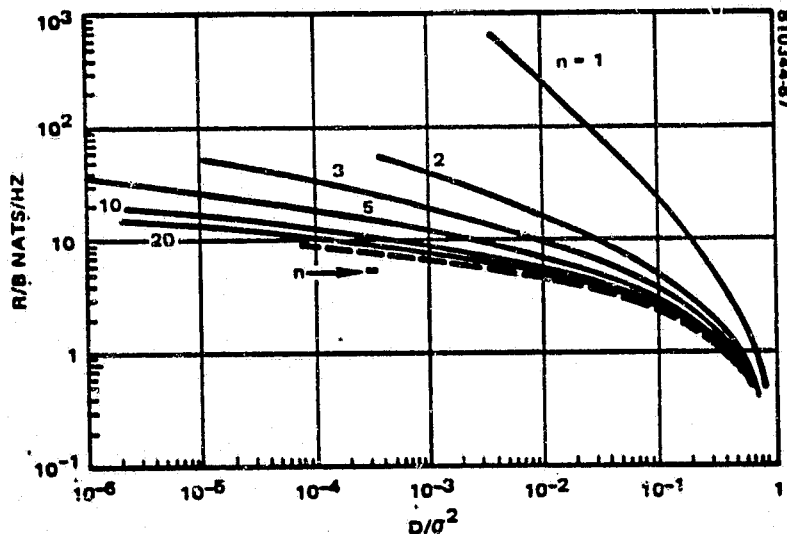
NOTE: GAUSSIAN MSE R (D) CURVES FOR $\phi(\omega) = A_n [1 + (\omega/\omega_0)^{2n}]^{-1}$; $\varepsilon [x^2 (t)] = \sigma^2$, $B = \omega_0/2\pi$.

FIGURE 10-4. BERGER'S PLOT OF R/B VERSUS D/σ^2

The corresponding function for the two-dimensional case is

$$R_2(D) = \frac{1}{2} \log_2 \left[\frac{\sigma^2 (1 - \rho_R^2)(1 - \rho_c^2)}{D} \right]$$

(I-10-11)

Figure 10-5 reproduces Pratt's Figure 7.7.2. There Pratt plots distortion functions for particular values of ρ_R and ρ_c.
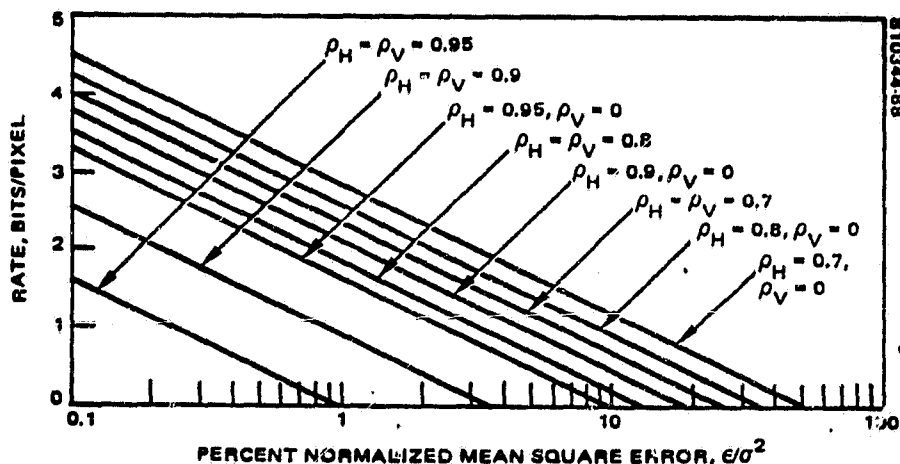

FIGURE 10-5. RATE DISTORTION FUNCTIONS FOR ONE AND TWO DIMENSIONAL CODING OF MARKOV IMAGE SOURCE

10.4 APPLICATION OF RATE-DISTORTION THEORY TO CALCULATION OF MFASMR PERFORMANCE

In Appendix I the MSE distortion is calculated for the $\omega^{-3/2}$ spectrum which we have selected as the spectrum of microwave radiometer scenes. Again both one- and two-dimensional cases are treated.

The cases treated in Section 10.2 are for signals which are continuous in amplitude. If the MSE distortion D becomes zero, the information rate becomes infinite. However, MFASMR will use digitized signals quantitized in amplitude. This limits information rate even if there is no distortion.

Figure 10-6 shows that good quantizers turn in a performance not much worse than the classical Shannon rate-distortion curve.

Not only is the MFASMR image signal not continuous in amplitude, it is also not spatially continuous. The image is characterized by the number of raster
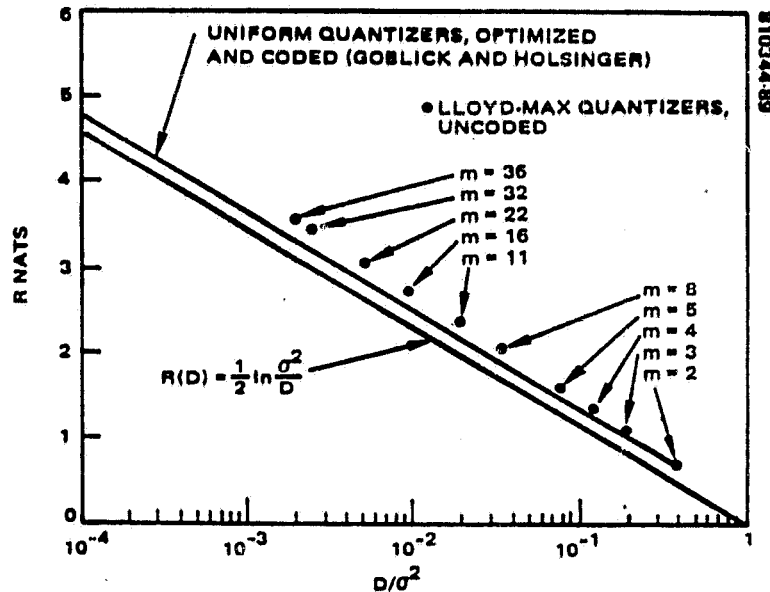
FIGURE 10-6. MAXIMUM INFORMATION RATE VERSUS
NUMBER OF QUANTIZATION BITS

lines present. This also limits information rate. In order to separate rate limitations due to the finite number of image lines and limitations due to undersampling, the spectrum used in Appendix I was normalized. The number of image lines was selected as an image parameter. Then we can write

$$\omega_{max} = M\omega_{min} \qquad\qquad (I\text{-}10\text{-}12)$$

The maximum information rate versus number of image lines was then calculated. The result is shown here as Figure 10-7.

The information rate was then normalized as a percent of the maximum possible rate for that number of raster lines. The distortion was normalized as a percent of the variance. The resulting rate distortion curves are plotted with the number of raster lines as a parameter. The result is shown as Figure 10-8.

The general region in which MFASMR systems will operate is indicated on Figure 10-8 as an oval centered on an information rate of 20 percent of the maximum rate obtained with Nyquist sampling. It may be seen that the mean-square distortion will be about one-third of the total signal power.
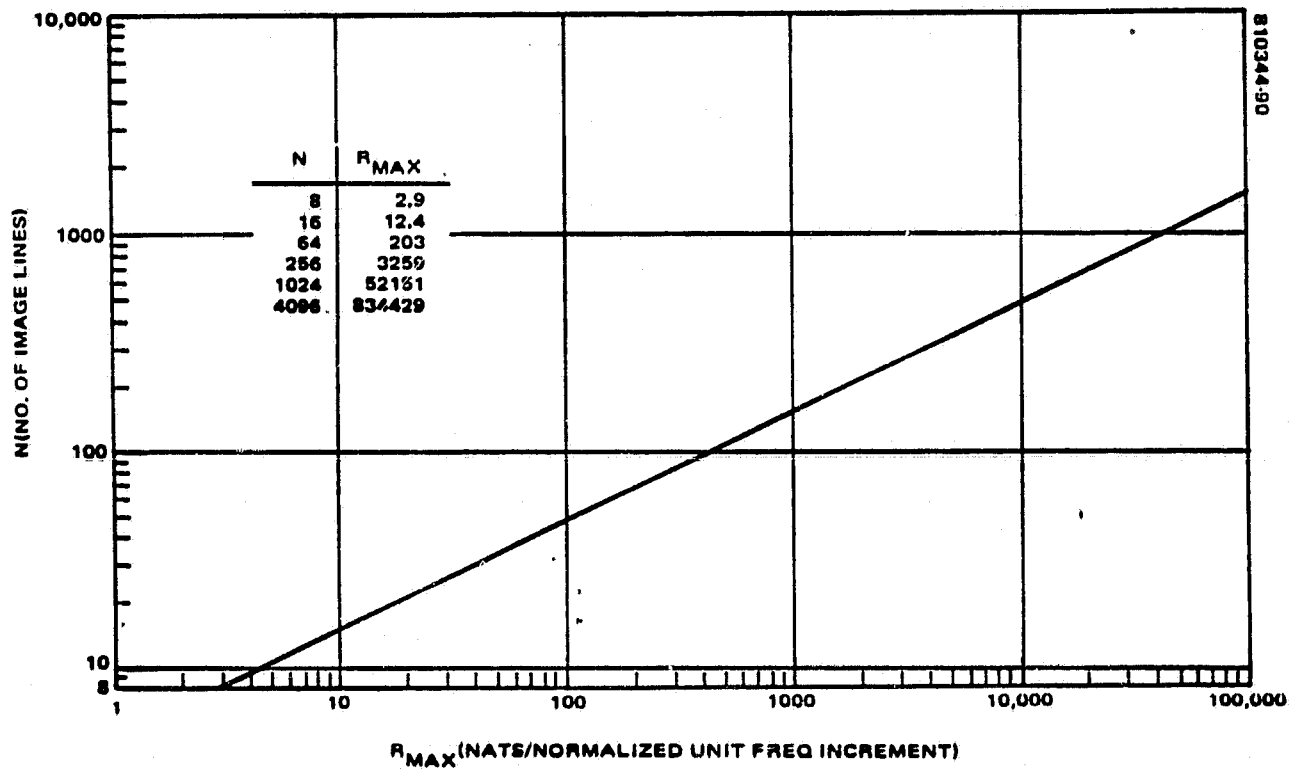
FIGURE 10-7. MAXIMUM INFORMATION RATE VERSUS NUMBER OF IMAGE LINES
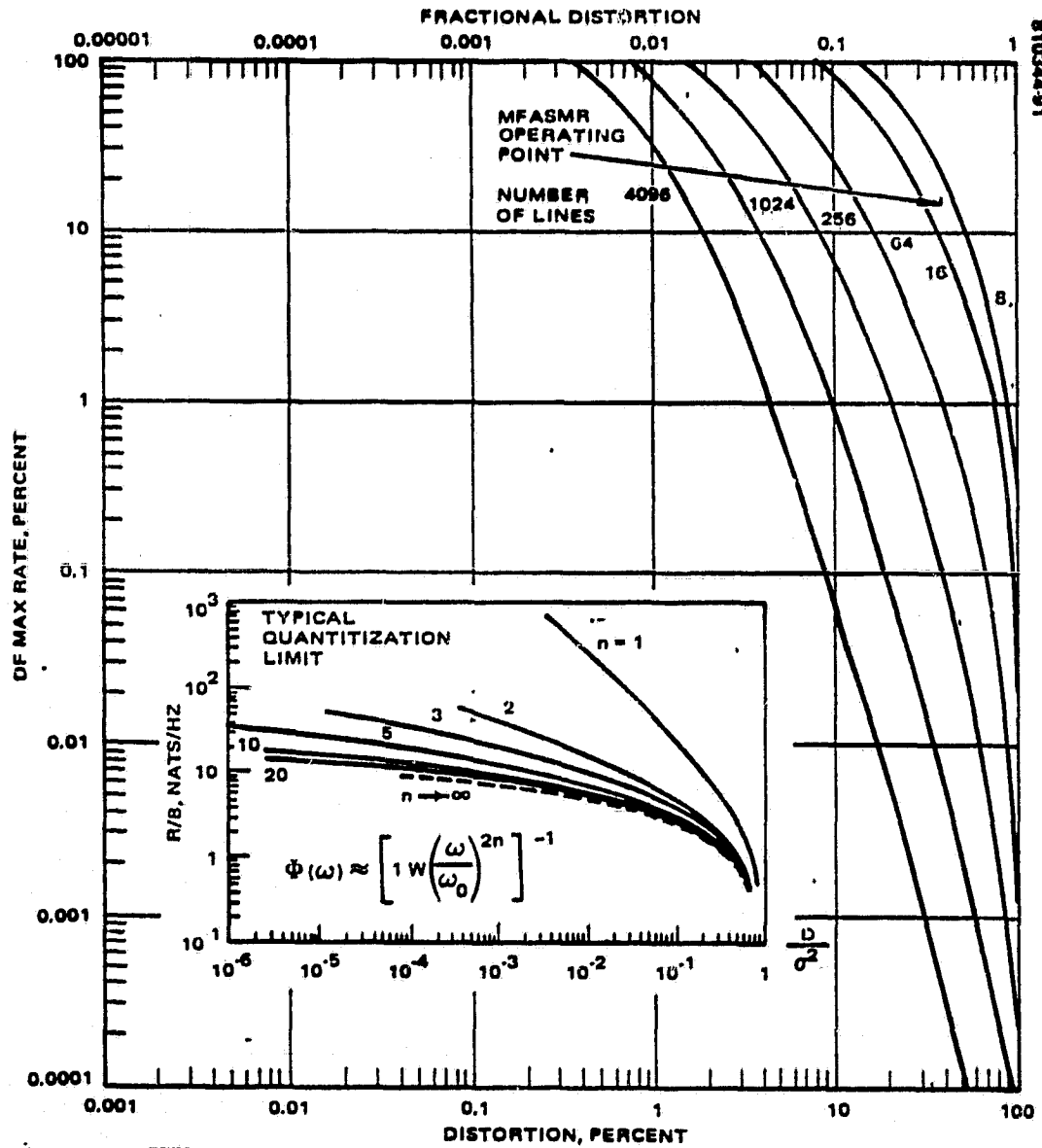
FIGURE 10-8. RATE-DISTORTION FUNCTIONS FOR $\Phi(\omega) = \omega^{-5/2}$

10.5 COMMENT ON MFASMR RATE DISTORTION CALCULATIONS

The reaction to the calculations in Section 10.3 is one of, "What went wrong?" In the original Hughes proposal in Figure 2-6 a comparison image distortion versus image compression ratio. This ratio is the reciprocal of the fractional undersampling. A compression ratio of 3:1 was estimated at that time and a distortion of 0.1 percent was read from the graph. This figure is reproduced here as Figure 10-9. We now know that the compression ratio is about 5 or 6 corresponding to the information rate of 20 percent of maximum rate. The lower result is due to the fact that we did not count samples from baselines which are translated, but not rotated, in this report. We now know these samples are redundant.

Figure 10-9 shows that for the less esoteric coding schemes and about 5:1 compression we should expect about 1 percent distortion. Now. Figure 10-9 is data obtained with optical images. Such images usually have 1,000 lines or so. Figure 10-8 shows that allowable data compression is a strong function of the number of image lines. In Figure 10-8 we see that 20 percent undersampling gives about 2.5 percent distortion when the image has a thousand lines. This is in rough agreement with Figure 10-9. If we could increase the number of lines in the MFASMR image to 1,000 or so, image distortion would be reasonable.
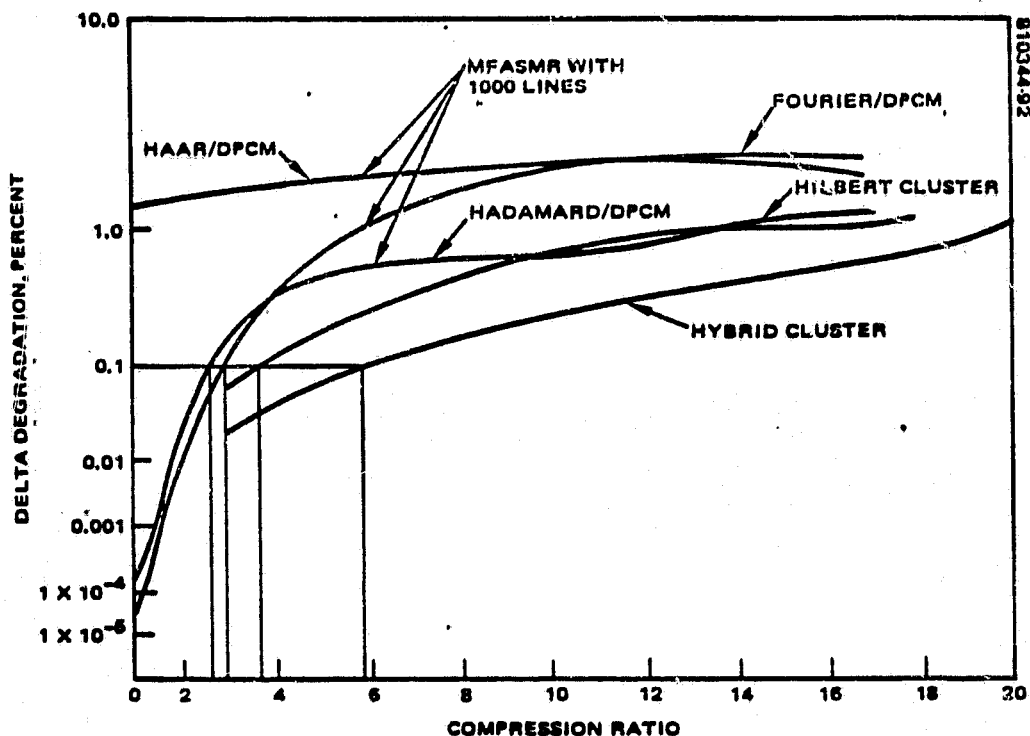
FIGURE 10-9. DISTORTION VERSUS COMPRESSION RATIO FOR VARIOUS CODING ALGORITHMS

Unfortunately there are not enough frequency bands available for use by radiometers to allow us to shrink the antennas and increase the number of lines while maintaining an information rate which is a fixed percentage of maximum rate.

The reluctant conclusion is that other frequency-domain imager designs will be more satisfactory.

Hopefully in the process of understanding MFASMR, we have arrived at a point where our understanding of frequency-domain imagers is complete enough to produce more successful designs.

We see that the probe antennas must be very small compared to the base-line length. This makes the signal record long enough to measure the spatial frequency spectrum with good accuracy. This in turn leads to a high resolution image after Fourier transformation of the record.

This is completely analogous to the SAR radar case. The SAR radar is an active frequency-plane imager. Just as in the passive case, the resolution of the SAR improves as the antenna size decreases.

As we have seen in earlier chapters satisfactory frequency-imaging systems for orbital use can be designed which do not use undersampling. Perhaps then we should start out by developing such systems. This is consistent with the history of radio-telescopes. The early instruments did not undersample. Then we can decide whether undersampling is beneficial in orbital mappers after exhausting the possibilities of systems meeting the Nyquist criterion.

11. BIBLIOGRAPHY FOR PART I

(1) S. J. Wernecke and L. R. D'Addario, "Maximum Entropy Image Reconstruction," _IEEE Transactions on Computers_, vol. C-26, No. 4, pp. 351-364, April 1977.

(2) S. J. Wernecke, "Two-Dimensional Maximum Entropy Reconstruction of Radio Brightness," _Radio Science_, vol. 12, No. 5, pp. 831-844, Sept.-Oct. 1977.

(3) S. J. Wernecke, "Maximum Entropy Techniques for Image Reconstruction from Interferometer Measurements and Projections," Ph.D Dissertation, _Stanford Electronics Laboratories_, Stanford, Ca., May 1976.

(4) S. J. Wernecke and C. J. Grebenkemper, "An Iterative Program for Maximum Entropy Fourier Synthesis," Stanford Radio Astronomy Institute Technical Memo #582, Stanford, Ca., 7 Dec. 1976.

(5) B. Oliver, "Thermal and Quantum Noise," pp 436-454, _Proc IEEE_, vol. 53 (May 1965); also pp 129-148 of IEEE Reprint Series Book, _Electrical Noise; Fundamentals and Sources_, Edited by M. Gupta. Proof is in the section entitled, "An Antenna Theorem and its Consequences."

(6) A. Hewish, "The Realization of Giant Radio Telescopes by Synthesis Techniques," pp 225-231, Sec 2(ii) _The Detection Sensitivity_, _Proc. I.R.E._ Australia, Feb 1963.

(7) M. Ryle and A. Hewish, M.N., 120, 220, 1960.

(8) J. Kraus, _Radio Astronomy_, Sec 6-10 and 6-11, pp 172-185, McGraw-Hill, 1966.

(9) T. Otoshi, "The Effect of Mismatched Components on Microwave Noise-Temperature Calibrations," _IEEE Trans on MTT_, Vol MTT-16, No. 9, Sept 1968.

(10) R. Bracewell et al, "The Stanford Five-Element Radio Telescope," _Proc IEEE_, Vol 61, No. 9, pp 1249-1257, Sept 1973.

(11) J. Baars et al, "The Synthesis Radio Telescope at Westerbork," pp 1258-1266, Proc IEEE, Sept 1973.

(12) K. Wesseling, "A Single-Sideband-Double-Sideband Interferometer Receiver for Radio Astronomy," IEEE Trans. on Antennas and Propagation, pp 332-333, March 1967.

(13) T. Seling, "The Application of Automatic Gain Control to Microwave Radiometers," pp 636-639, IEEE Trans. on Antennas and Propagation, Sept 1964.

(14) W. Groggins, Jr., "A Microwave Feedback Radiometer," pp 83-90, IEEE Trans. on Aerospace and Electronic Systems.

(15) J. Hach, "A Sensitive Airbourne Microwave Radiometer Using Two Reference Temperatures," pp 629-635, IEEE Trans., Vol MTT-16, No. 9, Sept 1968.

(16) W. Hardy et al, "An S-Band Radiometer Design with High Absolute Precision," pp 382-390, IEEE Trans on MTT, Apr 1974.

(17) A. Love and M. Van Meile, A Satellite Prototype S-Band Radiometer for Ocean Surface Temperature Measurement, Final Report, North Amer. Rockwell Space Div., Doc. No. SD72-SA-0119, 9 June 1972.

(18) M. Hersman and G. Poe, "Sensitivity of the Total Power Radiometer with Periodic Absolute Calibration," 16 Jan 1980, to be published in MTT Trans.

(19) M. Meeks (ed.), Astrophysics, Vol 12, Part B; Radio Telescopes, pp 266-277, Sec. 3.4 "Multichannel Spectrometers" by Hays Penfield, Academic Press, 1976.

(20) B. Lathi, Signals, Systems and Communication, Sec. 12.3, "Some Properties of Correlation Functions, pp 519-524, Wiley & Sons, 1965.

(21) M. Long and J. Butterworth, "A New Technique for Microwave Radiometry," pp 389-396, IEEE Trans. on MTT, Sept 1963.

(22) S. Weinreb, A Digital Spectral Analysis Technique and Its Application to Radio Astronomy, Mass Inst. of Tech. Research Laboratory of Electronics Tech. Report 412, 30 Aug 1963.

(23) F. Bowers et al, "A Digital Correlation Spectrometer Employing Multiple-Level Quantitization," pp 1339-1343, Proc IEEE, Vol 61, No. 9, Sept 1973.

(24) M. Meeks (ed), Astrophysics Part B: Radio Telescopes, Vol 12, Chapter 3.5, "Autocorrelation Spectrometers" by B. Cooper, pp 280-298, Academic Press, 1976.

(25) C. Papas, Theory of Electromagnetic Wave Propagation, pp 159-168, Sec 5.8, "Correlation Interferometer," McGraw-Hill, 1965.

(26) K. Fujimoto, on the Correlation Radiometer Technique, pp 203-211, Trans. on Microwave Theory and Technique, Vol MTT-12, March 1964.

(27) M. Cohen, "Introduction to Very-Long-Baseline Interferometry," pp 1192-1197, Proc IEEE, Vol 61, No. 9, Sept 1973.

(28) M. Born and E. Wolf, Principles of Optics, pp 554-555, Sec 10.8.3, "The Stokes Parameters of a Quasi-Monochromatic Plane Wave," Pergamon Press, 1964.

(29) H. Ko, "Theory of Tensor Aperture Synthesis," IEEE Trans on Antennas and Propagation, pp 188-189, Vol AP-15, No. 1, Jan 1967.

(30) H. Ko, "Coherence Theory of Radio Astronomical Measurements," pp 10-20, IEEE Trans on Antennas and Propagation, Vol AP-15, No. 1, Jan 1967.

(31) G. Evans and C. McLeish, RF Radiometer Handbook, pp 17-34, Chapter Two "Radiometer Systems," pp 105-112, Appendix A "Radiometer Sensitivity," Artech House, 1977.

(32) S. Goldstein, Jr., pp 1663-1666, "A Comparison of Two Radiometer Circuits," Proc IRE, Nov 1955.

(33) M. Tiuri, "Radio Astronomy Receivers," pp 930-938, IEEE Trans on Antennas and Propagation, AP-12, Dec 1964.

(34) J. Schuchardt and J. Stratigos, "Detected Noise Levels Guide Radiometer Design," pp 64-74, Microwaves, Sept 1978.

(35) J. Winderman, "Perform True DSB-to-SSB Noise-Figure Conversions," pg 69, Microwaves, July 1980.

(36) J. Bendat and A. Piersol, Measurement and Analysis of Random Data, pg 278, Wiley & Sons.

(37) R. Bracewell, The Fourier Transform and Its Applications, pp 337-340, Measurement of Noise Power, Wiley & Sons, 1978.

(38) J. Faris, "Sensitivity of a Correlation Radiometer," pp 153-170, Jour. of Research of the Nat. Bureau of Standards, Vol 71C, Apr - June 1967.

(39) W. Lindsey, Synchronization Systems in Communication and Control, pp 76-82, Sec 3.5, Prentice-Hall, 1972.

(40) K. Wesseling, "A Single-Sideband-Double-Sideband Interferometer Receiver for Radio Astronomy," IEEE Trans. AP-15, No. 2, pp 332-333, March 1967.

(41) M. Ryle, "A New Radio Interferometer and Its Application to the Observation of Weak Radio Stars," _Proc. Roy. Soc._ (London), Vol 211, pp 351-375, March 1952.

(42) S. Weinreb, "Low-Noise Cooled GASFET Amplifiers," _IEEE Trans._ Vol MMT-28, No. 10, pp 1041-1054, Oct 1980.

(43) U. Cardiasmenos, "Planar Devices Make Production Practical," _Microwave System News_, pp 15-21, April 1979.

(44) R. Clapp and J. Maxwell, "Complex-Correlation Radiometer," _IEEE Trans._, Vol AP-15, No. 2, pp 286-290, March 1967.

(45) J. Hubbard and W. Erickson, "A Stabilized Cross-Correlation Radiometer," _IEEE Trans._, Vol AP-15, No. 2, pp 291-294, March 1967.

(46) J. Lawson and G. Thlenbeck, _Threshold Signals_, MIT Rad. Lab. Series Vol 24, Sec 4.6 - 4.8, pp 60-63, McGraw-Hill, 1950.

(47) MSC Corp, _Data Sheet NM-101A and 101, Application Note NM-201_, Company address: 100 School House Rd., Somerset, N. J. 08873.

(48) .B. Mandelbrot, _Fractals-Form, Chance and Dimension_, W. Freeman and Co., 1977.

(49)

(50) A. J. Viterbi and J. K. Omura, _Principles of Digital Communication and Coding_, Sec. 8.4.3 "Continuous Time Sources and Generalizations," pp 502-513, McGraw-Hill, New York, 1979.

(51) T. Berger, _Rate Distortion Theory_, Sec. 4-5, p 123, Prentice-Hall, New Jersey, 1971.

(52) W. Pratt, _Digital Image Processing_, Sec 7.7, "Rate Distortion Function Representation," pp 189-192, Wiley & Sons, 1978.

(53) E. O'Neill, _Introduction to Statistical Optics_, Chapter 2, "Spatial Versus Time Filters," pp 12-29, Addison-Wesley, 1963.

(54) M. Schwartz and L. Shaw, _Signal Processing:_, pp 168-178, Sec 4.5 "Periodogram Revisited and Modified."

(55) G. Jenkins and D. Watts, _Spectral Analysis and Its Applications_, Holden-Day, San Francisco, 1968.

(56) W. Price, "The Photographic Lens," pp 72-83, _Scientific American_, Vol 235, No. 2, Aug 1976.

(57) F. Yu, _Optics and Information Theory_, Sec 3.1 "Optical Spatial Communication channel," pp 62-67, Wiley & Sons, 1876

(58) H. Booker and P. Clemmow, "The Concept of an Angular Spectrum of Plane Waves and Its Relation to That of Polar Diagram and Aperture Distribution," Proc. Inst. Elec. Engrs., London, ser. 3, Vol 97, pp 11-17, Jan 1950.

(59) J. Kraus, Radio Astronomy, Sec 6-9, "Spatial Frequency Response and Pattern Smoothing," pp 169-171, McGraw-Hill, 1966.

(60) E. Brigham, The Fast Fourier Transform, Sec 4.9, "Correlation Theorem," pp 66-68, Prentice-Hall, 1974.

(61) F. Hagen, The Effects of Aliasing on the SSMI Radiometer, Appendix A, Internal HAC Space/Comm Rpt, 23 Nov 1979.

(62) C. Drane, Jr. and J. McIlvenna, Response Function for Taylor Antenna Distributions, Air Force Cambridge Research Laboratories Rpt. AFCRL-63-368 and Defense Doc. Ctr. No. AD435621, Sept 1963.

(63) R. Utukuri and R. MacPhie, "Coincident Arrays for the Direct Measurement of the Principal Solution in Radio Astronomy," IEEE Trans. on Antennas and Propagation, Vol AP-15, No. 1, Jan 1967.

(64) J. Goodman, Introduction to Fourier Optics, McGraw-Hill, 1968.

(65) A. Papoulis, Systems and Transforms with Applications in Optics, McGraw-Hill, 1968.

(66) M. Born and E. Wolf, Principles of Optics, pp 508-510, Sec 10.4.2 "Calculation of Mutual Intensity and Degree of Coherence for Light From an Extended Incoherent Quasi-Monochromatic Source," Pergamon Press, 1964.

(67) R. Bracewell, The Fourier Transform and Its Applications, Chapter 8, pp 148-153, "Equivalent Width," McGraw-Hill, 1978.

(68) H. Ko, "Coherence Theory in Radio-Astronomical Measurements," IEEE Trans. Antennas and Propagation, AP-15, No. 1, pp 10-20, Jan 1967.

(69) B. Oliver, "Thermal and Quantum Noise," Proc IEEE, Vol 53, pp 436-454, May 1965. Also reprinted in Electrical Noise: Fundamentals and Sources, M. Gupta (ed), pp 129-143, IEEE Press.

(70) G. Swenson, Jr. and N. Mathur, The Interferometer in Radio Astronomy, Proc IEEE, Vol 56, No. 12, pp 2114-2129, Dec 1968.

(71) E. Fomalont, "Earth-Rotation Aperture Synthesis," Proc IEEE, Vol 61, No. 9, pp 1211-1218, Sept 1973.

(72) Steel, Interferometry, pg 222, Fig 91, Cambridge Press, 1967.

(73) A. Papoulis, <u>Signal Analysis</u>, Sec 10-2, Unknown Signals in Noise, pp 329-335, Fig 10-5, McGraw-Hill, 1977.

(74) A. Oppenheim and R. Shafer, <u>Digital Signal Processing</u>, Sec 11.3.3 General Variance Expressions, pp 545-548, Fig 11.3, Prentice-Hall, 1975.

(75) M. Harwit and N. Sloane, <u>Hadamard Transform Optics</u>, Academic Press, 1979.

(76) J. Bendat and A. Piersol, <u>Measurement and Analysis of Random Data</u>, Sec 5.3.5 Spectral Density Estimates, pp 195-200, Eq. (5.95), Wiley & Sons, 1966.

(77) R. Bracewell, <u>The Fourier Transform and Its Applications</u>, pp 251-253, McGraw-Hill, 1968.

(78) K. Sheridan, Techniques for the Investigation of Solar Radio Bursts at Metre Wavelengths, Proc I.R.E Australia, Vol 24, pp 174-184, Feb 1963.

(79) R. Johnson, Optical Scanners, Fig 20, pg 234 of <u>Microwave Scanning Antennas</u>, Vol I, R. Hansen (ed), Academic Press, 1964.

(80) J. Foster, A Microwave Antenna with Rapid Sawtooth Scan, Can. J. Phy. 36, pp 1652-, 1958.

(81) R. Honey and E. Jones, A Mechanically Simple Foster Scanner, IRE Trans. Ant. and Prop.

(82) G. Vanesse and H. Saki, Fourier Spectroscopy, pp 261-327, <u>Advances in Optics VI</u>, Wolf (ed), 1967.

(83) H. Yoshinaga, II Recent Developments in Far Infrared Optics, pp 79-121, <u>Advances in Optics XI</u>, Wolf (ed).

(84) C. Jackson, The Allocation of the Radio Spectrum, pp 34-39, <u>Sci. Amer.</u>, Vol 242, No. 2, Feb 1980.

(85) R. Bracewell, "Strip Integration in Radio Astronomy," <u>Australian J. Phys.</u>, Vol 9, pp 198-197, 1956.

(86) R. Bracewell and A. Riddle, "Inversion of Fan-Beam Scans in Radio Astronomy," <u>Astrophys. Jour.</u>, Vol 150, pp 427-434, Nov 1967.

(87) J. Taylor, "Brightness Distributions From Lunar Orcultations," pp 421-426, Vol 150, Nov 1967.

(88) A. Thompson and R. Bracewell, "Interpolation and the Fourier Transformation of Fringe Visibilities," pp 11-24, <u>The Astronomical Journal</u>, Vol 79, No. 1, Jan 1974.

(89) H. Stark, "Sampling Theorems in Polar Coordinates," pp 1519-1525,
J. Opt. Soc. Am., Vol 69, No. 11, Nov 1979.

(90) I. Cindrich et al, "Optical Processing Architecture for Radio
Telescope Data," pp 50-62, Vol 214 Acousto-optic Bulk Wave Devices,
Proc. Soc. of Photo-optical Instrumentation Engineers, Nov 27-29, 1979,
Monterey, Ca.

(91) I. Cindrich et al, "Optical Processor System for VLA Radio Telescope
Data," ERIM Report 123400-8-F, Apr 1977.

(92) R. Mitchel, SAR Image Quality Analysis Model, Thesis Dissertation for
E. E. Dept. Un. of Mich. (1974), Also ERIM Report (July 1974), Sec 6.2
"Holographic Viewer Concept."

(93) G. Parrent, J. Opt. Soc. Am. 49, pg 787, 1956.

(94) M. Beran and .G. Parrent, Jr., Theory of Partial Coherence, Sec 3.1
"Derivation of the Differential Equations Governing $\Gamma_{12}(\tau)$," pp 36-
38, Prentice-Hall, 1964.

(95) P. Karr, Generalized "Cross-Correlation" Field Quantities as Solutions
of Field Equations, Space Technology Labs, Inc Rpt 9990-6218-RV-000, STL
address: One Space Park, Redondo Beach, Ca., Aug 1962.

(96) L. Levi, Applied Optics, pp 76-77, Sec 2.3.1.2 "Diffraction by an
Arbitrary Aperture," Wiley & Sons, 1976.

(97) B. Steinberg, Principles of Aperture and Array System Design, pp 9-11,
Sec 1.4 "Huygen's Principle," Wiley & Sons, 1976.

(98) P. Beckman, Orthogonal Polynominals for Engineers, Sec 2.4 and Sec 3.1,
pp 23-33, Golem Press, 1973.

(99) A. Papoulis, Systems and Transforms with Applications in Optics,
Chap. 9, Sec 5, Eq 5-5, pp 345-354, McGraw-Hill, 1968.

(100) S. Wernecke, Maximum Entropy Techniques for Image Reconstruction from
Interferometer Measurement and Projections, Doctoral Dissertation
submitted to Stanford University E. E. Dept., May 1976.

(101) R. Crowther et al, "The Reconstruction of a Three-dimensional Structure
from Projections and Its Application to Electron Microscopy," Proc.
Roy. Soc. (London), ser. A, Vol 317, pp 319-340, 1970.

(102) B. Vainshtein, "The Synthesis of Projecting Functions," Sov. Phy.-
Doklady, vol 16, pp 66-69, 1971.

(103) R. Bates and T. Peters, "Towards Improvements in Tomography," New
Zealand J. Sci., Vol 14, pp 883-896, 1971.

(104) R. Gordon et al, "Algebraic Reconstruction Techniques (ART) for Three-dimensional Electron Microscopy and X-Ray Photography," _J. Theor. Biol._, Vol 29, pp 471-481, 1970.

(105) R. Gordon, "A Tutorial on ART," _IEEE Trans. Nucl. Sci._, Vol NS-21, pp 78-93, 1974.

(106) P. Gilbert, "Iterative Methods for the Reconstruction of Three-Dimensional Objects from Projections," _J. Theor. Biol._, Vol 36, pp 105-117, 1972.

(107) G. Herman, Chap. 10, "Convolution Methods for Divergent Beams," _Image Reconstruction from Projections_, Academic Press, 1980.

(108) C. Wiley, Chapter 8, "Airbourne Radar Navigation," Sec 8.6.2, "Unfocused SAR Systems," pp 361-363, _Avionics Navigation Systems_, (M. Kayton and W. Fried (ed), Wiley & Sons, 1969.

(109) M. Seiler, "Radio Interferometry of Random Terrestrial Backgrounds," pp 638-644, _IEEE Trans. Aerospace and Electronic Sys._, AES-5, No. 4, July 1969.

(110) Naval Air Systems Command Rpts No. M000-19-(67-C-0477) or (68-C-0390) or Final Rpt. (69-C-0324), _Radiometric Signal Processing and Display Techniques_, 1969.

(111) Rockwell International Final Rpt. No. MR-69H-502, 1969.

(112) C. Cook and M. Bernfeld, _Radar Signals_, Table 4-1, pp 88-91, "Ambiguity Function Properties," Fig 4-14, pg 96, "Parabolic FM Waveform and Response Functions," Sec 9.2, pg 304, "Minimum Time and Frequency Error Variances," Table 9-1, pg 307, "Accuracy Parameters of Unidirectional FM Pulse-Compression Signals, B. Parabolic FM: rect (t/τ) envelope," Sec 9.3, pg 314, "Effect of Range-Doppler Coupling on Measurement Errors, Parabolic FM (Unidirectional)," Academic Press, 1967.

(113) W. Rotman and R. Turner, "Wide-Angle Microwave Lens for Line Source Applications," pp 623-632, _IEEE Trans. Ant. and Prop._ 11, Nov 1963.

(114) R. Honey and M. Jones, "A Mechanically Simple Foster Scanner," pp 40-46, _IRE Trans. on Ant. and Prop._ 4, Jan 1956.

(115) D. Margerum, Chapter 5, "Self-Phased Arrays," pp 341-407, Vol III _Microwave Scanning Antennas_ (R. Hansen, ed), Academic Press, 1966.

(116) R. Monzingo and Miller, _Introduction to Adaptive Arrays_, Wiley & Sons, 1980.

(117) K. Castleman, Chap. 4, "Image Processing Software and Appendix II VICAR Program Index," _Digital Image Processing_, Prentice-Hall, 1979.

(118) B. Hunt, "Digital Image Processing," pp 196-211, from Digital Image Processing for Remote Sensing (R. Bernstein, ed), IEEE Press, 1978.

(119) T. Stockham, "Image Processing in the Context of a Visual Model," pp 212-226, Same volume as Ref. 118 above.

(120) J. Bendat, Sec 4.4, "Wiener Filter Problems," pp 153-147, Eq 4-17, Principles and Applications of Random Noise Theory, Wiley & Sons, 1958.

(121) D. Childers (ed), Modern Spectrum Analysis, IEEE Press, 1978.

(122) J. Ables, "Maximum Entropy Spectral Analysis," pp 32-33, Figs 1 to 4, Modern Spectrum Analysis.

(123) W. King, Maximum Entropy Spectral Analysis in the Spatial Domain, Rome Air Development Center Report RADC-TR-78-160, Figs 4, July 1978.

(124) S. Werneke, MEM Techniques for Image Reconstruction, Doctoral Dissertation submitted to Stanford Un. Elect. Eng. Dept., May 1976.

(125) S. J. Wernecke and C. J. Grebenkemper, "An Iterative Program for Maximum Entropy Fourier Synthesis," Stanford Radio Astronomy Institute Technical Memo #582, Stanford, Ca., 7 Dec 1976. NOTE: This document contains the MEM program listing. MPTX has been compiled by the FORTRAN II-extended compiler (OPT 2).

(126) S. J. Wernecke, "Two-Dimensional Maximum Entropy Reconstruction of Radio Brightness," Radio Science, vol. 12, No. 5, pp. 831-844, Sept-Oct. 1977.

(127) S. J. Wernecke and L. R. D'Addario, "Maximum Entropy Image Reconstruction," IEEE Transactions on Computers, vol. C-26, No. 4, pp. 351-364, April 1977.

(128) K. Castelman, pp 283-288, Digital Image Processing, Prentice-Hall, Inc., 1979.

(129) S. Goldman, Sec 4.9, "Entropy Loss in Filters," pp 144-147, Information Theory, Prentice-Hall, 1953.

(130) J. Ohlson, "On the Optimum Radiometer," Radio Science, Vol 8, Number 10, pp 841-844, Oct 1973.

(131) A. Sawchuk, "Space-Variant Image Motion Degradations and Restoration," Proc IEEE, Vol 60, No. 7, pp 854-861, July 1972.

(132) A. Sawchuk, "Space-Variant System Analysis of Image Motion," JOSA, Vol 63, No 9, pp 1052-1063, Sept 1973.

(133) A. Sawchuk, "Space-Variant Image Restoration by Coordinate Transformation," JOSA, Vol 64, No. 2, pp 138-144, Feb 1974.

(134) F. Yu, Sec 7.6 "Restoration of Blurred Photographic Images,"
pp 161-168, <u>Optics and Information Theory</u>, Wiley & Sons, 1976.

(135) J. Tsujiuchi, "Correction of Optical Images by Compensation of Aberrations
and Spatial Frequency Filtering," in <u>Progress in Optics</u>, Vol II,
E. Wolf (ed), North-Holland, 1963.

(136) E. Jacobs, "On the Spatial Frequency Response of an Antenna,"
<u>Proc. IEEE Letters</u>, pg 92, Jan 1966.

(137) J. Bendat and A. Piersol, Sec 6.5 "Spectrum Density Functions Estimates,"
pg 187, Eq 6.95, <u>Random Data: Analysis and Measurement Procedures</u>,
Wiley & Sons, 1971.

(138) B. Steinberg, Chapter 11, "Adaptive Beamforming," pp 212-252,
<u>Principles of Aperture and Array System Design</u>, Wiley & Sons, 1976.

(139) C. Dorney and B. Meager, "Cohering of an Experimental Nonrigid Array
by Self-Survey," Trans. Ant. and Prop. Vol AP-28, No. 6, pp 902-904,
Nov 1980.

(140) T. Berger, <u>Rate Distortion Theory, A Mathematical Basis for Data
Compression</u>, Prentice-Hall, 1971.

(141) T. Berger, Ibid, pp 106-130, Sec 4.5, R(D) for Gaussian Sources.

(142) A. Viterbi and J. Omara, <u>Principles of Digital Communication and
Coding</u>, Sec 8.4.3, "Continuous Time Sources and Generalizations and
Coding," pp 502-513, McGraw-Hill, 1979.
(143) W. Pratt, Sec 7.7 "Rate Distortion Function Representation," pp 189-192,
<u>Digital Image Processing</u>, Wiley & Sons, 1978.

(144) A. K. Jain, "Image Data Compression: A Review," <u>Proc. of IEEE</u>,
Vol. 69, pp. 349-389, March 1981.

(145) J. R. Jain, A. K. Jain and R. A. Robb, "Data Compression of Multi-
dimensional X-ray Images," to be published.

PART II

Computer Study
of
Image Reconstruction with MFASMR Data

1. INTRODUCTION

1.1 PROBLEMS OF IMAGE RESTORATION

As presented in Part I, digital image restoration is a mandatory element in designing a frequency domain imaging system. This is especially true if the system considered corresponds to sampling below Nyquist rate.

In the conventional imaging system, the spatial/temporal images are formed directly by hardware. Depending on the application of the images, the removal of the modulation transfer function (MTF) of the antenna or the point spread function (PSF) of the sensor by means of image restoration may or may not be necessary. In the case of frequency domain imager, the spectral data must at least be Fourier transformed to produce the spatial/temporal images.

For the time being, assume that the imaging system is linear. Following the notations in Reference 1, let $f(x,y)$ be the imaging scene, and let $h(x,y,u,v)$ the system transfer function, then the sensor output g is obtained by

$$g(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x,y,u,v) f(u,v) \, du \, dv$$

or,

$$T\{f\} \rightarrow g \tag{II-1-1}$$

The problem is to find T^{-1} such that $T^{-1}\{g\} \rightarrow f$. For a fully sampled frequency imager, $h(x,y,u,v) = \exp[-j2\pi(xu+yv)]$. In this case, T corresponds to a two-dimensional Fourier transform, and T^{-1} is then the inverse transform.

In an undersampled system, a nonlinear sampling function $s(\cdot)$ is applied to g to produce measurements m_i, $i = 1,2,3, \ldots M$

$$m_i = s(g(x,y)) + n_i \tag{II-1-2}$$

when n_i is the measurement noise/error. To reproduce $f(x,y)$ from m_i is no longer a trivial task. Reference 1 pointed out that the problem of image restoration is an ill-conditioned problem at best and singular at worst. T^{-1} may not exist as in the case of under sampled system (singular) nor does it have to be unique. Even if T^{-1} exists and is unique, (as in the case of fully sampled frequency domain imager), it may be so ill-conditioned that for a small perturbation ε in the measurement field g,

$$T^{-1}\{g+\varepsilon\} \rightarrow f + d \qquad\qquad (II-1-3)$$

where d is not arbitrarily small and is not negligible. Chapter 6 of Reference 1 gives an excellent discussion of above problems.

1.2 RATE DISTORTION THEORY FOR DATA COMPRESSION

In cases where perfect reconstruction of the image is not required, i.e., under certain distortion constraints, rate distortion theory is the mathematical basis for data compression (Reference 2). It should be noted that data compression exploits the statistical properties of the information source and removes redundancy, subject to a distortion measure. In this way, useful information is not arbitrarily discarded.

Consider a discrete memoryless source X. It is desirable to encode X such that Y, an estimate of X, can be reconstructed. See Figure 1-1.

Let $p(x)$ be the source probability density and $d(x,y) \geq 0$ be a single-letter distortion measure. If $p(x,y)$ is the joint probability of X and Y, then

$$R(D) = \min I(X;Y) \qquad\qquad)II-1-4)$$

$$\{p(x,y): \sum_y p(x,y) = p(x) \text{ and } \sum_x \sum_y p(x,y)d(x,y) \leq D\}$$

where $I(X;Y)$ is the Shannon average mutual information between the two random processes X and Y (Reference 2). $R(D)$ represents the minimum information rate that an encoder-decoder pair can operate on X such that the average distortion between X and Y, an estimate of X, is no longer than D.
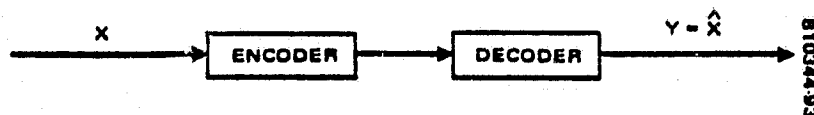
FIGURE 1-1. SOURCE CODING BLOCK DIAGRAM

It is easier to understand the mutual information if it is expressed by

$$I(X;Y) = H(X) - H(X|Y) \qquad (II-1-5)$$

where H(X) is the entropy of X, $H(X) = \sum_x p(x)\log(p(x))$; and H(X|Y) is the conditional entropy (i.e., the remaining uncertainty of X when Y is completely given). To encode-decode X noiselessly, Y is required to supply all the information about X; therefore H(X|Y) = 0 or given Y there is no more uncertainty in X. Equation (II-1-5) yields R(0) = H(X), which is the Shannon noiseless coding theorem; one needs the rate equal to the source entropy to code it noiselessly. If Y is such that X and Y are independent, then H(X|Y) = H(X). In this case, Y supplies no information about X and thus coding/decoding process suffers a maximum distortion. This can be illustrated by a simple example. Let X be a coin toss process with a fair coin and Y be an independent coin toss process. On the average, X and Y are going to have the same outcome (head or tail) half the time. If d(x,y) = 1 for x ≠ y and d(x,y) = 0 for x = y (Hamming distortion), then the average distortion is 0.5. Therefore no information about the X process has to be encoded. At the decoder end, one only has to perform an independent experiment which has the same probability distribution as X to achieve the maximum distortion of 0.5 (R(0.5) = 0.).

For a continuous amplitude stationary process (analog source), the entropy rate analogous to the discrete entropy goes to infinity. The extension of the rate distortion function to such sources is not trivial. It can be shown that for a zero mean discrete-time stationary Gaussian source with mean-square distortion measure (d(x,y) = (x-y)^2), R(D) can be parameterized by

$$R(D) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \max[0, \log \frac{\phi(w)}{\theta}] \, dw$$

$$D = \frac{1}{2\pi} \int_{-\pi}^{\pi} \min[\theta, \phi(w)] \, dw \qquad (II-1-6)$$

where $\phi(w)$ is the spectral density of the process; and

$$\sigma^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(w) \, dw \quad < \quad \infty \qquad (II-1-7)$$

is the variance of the process. Figure 1-2 shows few typical examples of R(D) curves for such Gaussian sources with various correlation properties (correspond to different spectral densities). These curves represent the optimum encoder/decoder pair that one can operate and achieve the desired distortion. Equation (II-1-6) can be extended to the time continuous case by taking the limit of the integrals to infinity. Figure 1-3 gives the graphical interpretation of Eq. (II-1-6). In ideal MSE econding of the Gaussian source, it is necessary to reproduce only the portion of the source output at frequencies where $\phi(w) > \theta$. If θ is such that only the centerlobe of $\phi(w)$ is above θ, then an ideal low pass filter is the only necessary component in the source encoder.
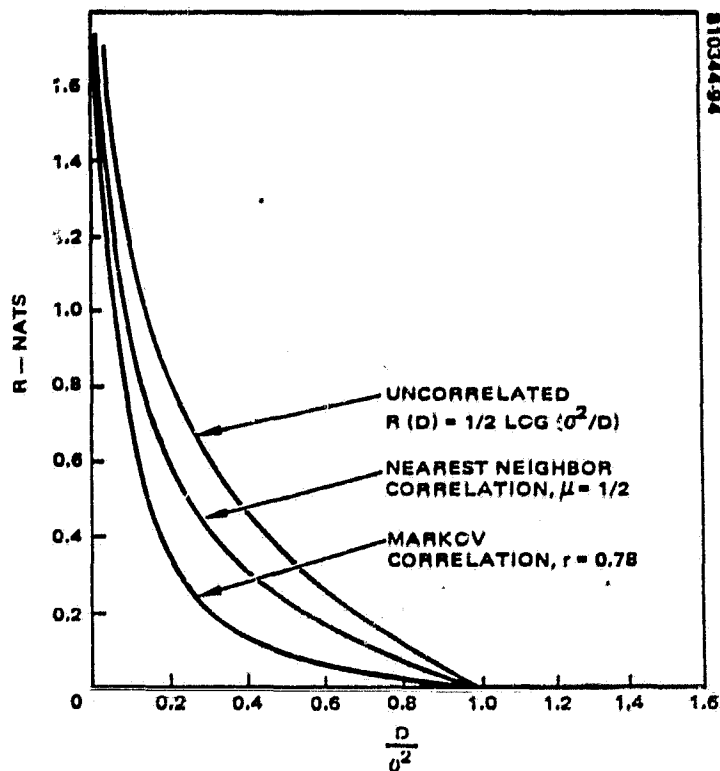
FIGURE 1-2. COMPARISON OF MSE RATE DISTORTION
FUNCTIONS FOR UNCORRELATED DATA
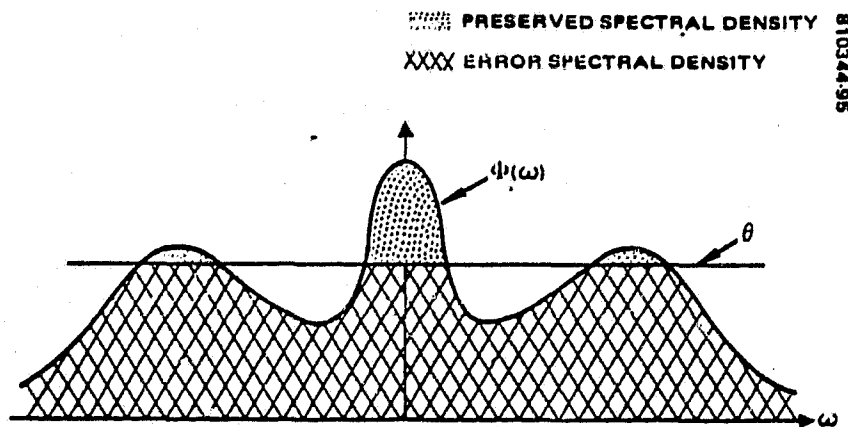
FIGURE 1-3. SPECTRAL DENSITIES OF SOURCE, ERROR,
AND REPRODUCTION FOR OPTIMUM CODING

Equation (II-1-6) can be extended to two dimensions in order to apply to Gaussian random fields (Reference 3),

$$R(D) = \frac{1}{8\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \max\left[0, \log \frac{\phi(u,v)}{\theta}\right] du\, dv$$

$$(II-1-8)$$

$$D = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \min[\theta, \phi(u,v)] du\, dv$$

where u and v are the spatial frequencies of the two-dimensional Gaussian random field. Equations (II-1-6) and (II-1-8) which define R(D) for Gaussian sources, are the upper bounds to all other rate distortion functions of non-Gaussian sources with the same spectral density ϕ. Part I uses the above formula applying $\phi(p) = p^{-5/2}$, $p = \sqrt{u^2+v^2}$, as the spectral density of the radiometer image. It gives an indication as the achievable distortion if the data are undersampled.

There are few restrictions when applying the theory. The source statistics have to be known although work has been done to extend the theory to universal source coding to apply to sources with completely or partially unknown statistics. More severely, the theory provides only the theoretical limitations of expected performance. It does not in any way provide the necessary algorithm to achieve this theoretical limitation. As suggested by the minimization problem of Eq. (II-1-4) over probability distributions, the mapping from X to Y is probabilistic and is governed by $p(x|y)$ in achieving the minimum. Often, one can apply Eq. (II-1-6) using the estimate of the one dimensional spectral density of the random process. This estimate may be obtained by average periodigram or maximum entropy techniques. Part I discussed some of the pros and cons of these techniques. However, in order to apply Eq. (II-1-8) to two-dimensional images, the two-dimensional power spectral density has to be estimated. Roucos and Childers (Reference 4) recently extended the maximum entropy spectral estimation method to the two-dimensional case but does not provide a guarantee of the convergence of solutions. Finally, a meaningful distortion criterion for subjective image quality has yet to be defined mathematically so that one can incorporate it in the minimization of Eq. (II-1-4).

Image data compression techniques have undergone extensive research (References 5 through 7). In every case, the redundancy of the images is carefully studied for its particular application. It is emphasized that the objective of data compression is to produce a mapping (or transformation) T, (which may not be linear), such that a suitable T^{-1} can be found to reconstruct the original images. Image restoration, however, usually deals with only the second half of the above problem to remove inherent system degradations.

1.3 FOURIER SYNTHESIS TEST CASES

In simulating a frequency domain imaging system, a spatial frequency plane is obtained from an image $f(x,y)$ via

$$F(u,v) = \iint_D h(x,y,u,v) f(x,y) \, dx \, dy \qquad \text{(II-1-9)}$$

where $h(x,y,u,v) = \exp[-j2\pi(xu+yv)]$ and D is the field of view. Note that since $f(x,y)$ is real, $F(u,v)$ is conjugated symmetrically. Measurements, m_i, have to be made only on half of the frequency plane:

$$m_i = S\{F(u,v)\} + e_i \qquad i = 1,2,3, \ldots M \qquad \text{(II-1-10)}$$

where S is the sampling function and e_i is the measurement error. Three sampling strategies are used in studying the reconstruction techniques:

$$S1\{F(u,v)\} = \begin{cases} F(u,v) & \text{if } u=v, \ u=-v, \ u=0, \ v=0 \\ 0 & \text{else} \end{cases} \qquad \text{(II-1-11)}$$

This is similar to multi-λ system's sampling of frequency axes and diagonals in Cartesian coordinate. This is an extremely undersampled system. For a 64x64 image, $S1\{F(u,v)\}$ only retains 249 samples from $F(u,v)$ over the spatial frequency domain coverage.

$$S2\{F(u,v)\} = \begin{cases} F(u,v) & \text{if } |u| \leq (1/2)u_{max}, \ |v| \leq (1/2)v_{max} \\ 0 & \text{else} \end{cases} \qquad \text{(II-1-12)}$$

This is similar to ideal lowpass filtering, only the frequency samples near zero spatial frequency are preserved.

$$S3\{f(u,v)\} = \begin{cases} F(u,v) & \text{if } |u| \leq (1/4)u_{max}, \ |v| \leq (1/4)v_{max} \\ 0 & \text{else} \end{cases} \qquad \text{(II-1-13)}$$

For a 64x64 image, this corresponds to low pass filtering with about same number of measurements as S1 (256 vs. 249).

Figures 1-4 and 1-5 illustrate the 64x64 stacked cube test patterns along with their two-dimensional Fourier transforms. Case 1 clearly has its spatial frequency contents concentrated along axes and diagonals. Case 2 is rotated slightly with respect to Case 1, in order to disperse its spatial frequency contents. Both cases are essentially the checker board test patterns.

Two basic reconstruction methods are investigated. The simplest one is the direct transform technique which includes inverse filtering and Wiener filtering. For singular reconstructions, inverse and Wiener filtering are not expected to provide any improvements. They are included in this report because of their classical role in image restoration. Examples are given to demonstrate that both methods result in an improvement over direct inverse transforming when the system is not undersampled. The second method, due to Wernecke (References 8 through 11), is iterative in nature and is called maximum entropy method (MEM). Brief descriptions of each method are described in their respective chapters. In Section 4, Part II, three radiometer images are used to study the reconstruction algorithms and their performances are compared. Thoughout the report, squared-errors are computed by taking the sum of the squares of the differences between original and reconstructed images. The sum is then divided by the total number of pixels to produce estimated mean-square-error.

CASE 1
ORIGINAL

810344-96
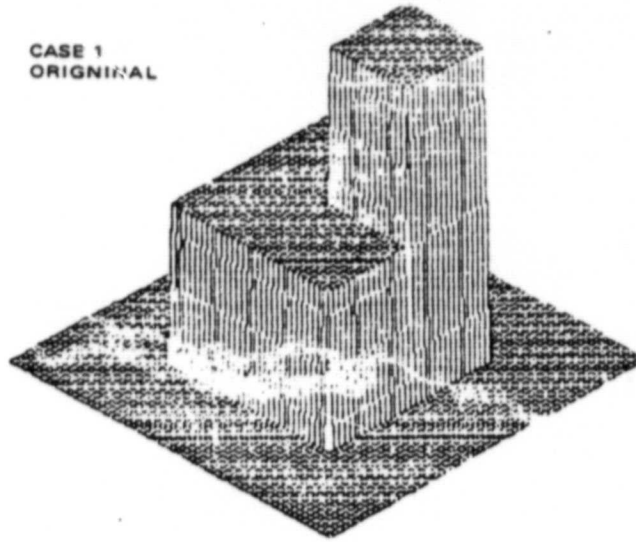
CASE 1
REAL PART OF
FOURIER TRANSFORM
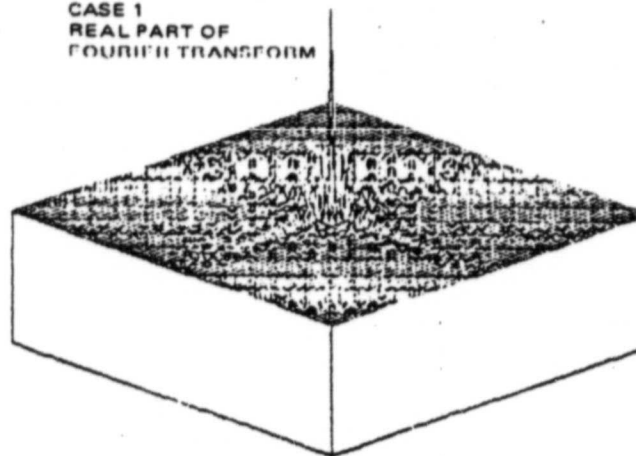
CASE 1
IMAGINARY PART OF
FOURIER TRANSFORM
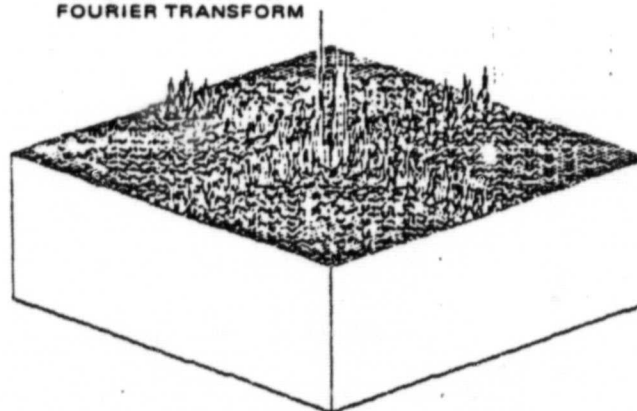
FIGURE 1-4. STACKED-CUBE TEST CASE 1

II-1-8

810344-97

CASE 2
ORIGINAL

a)
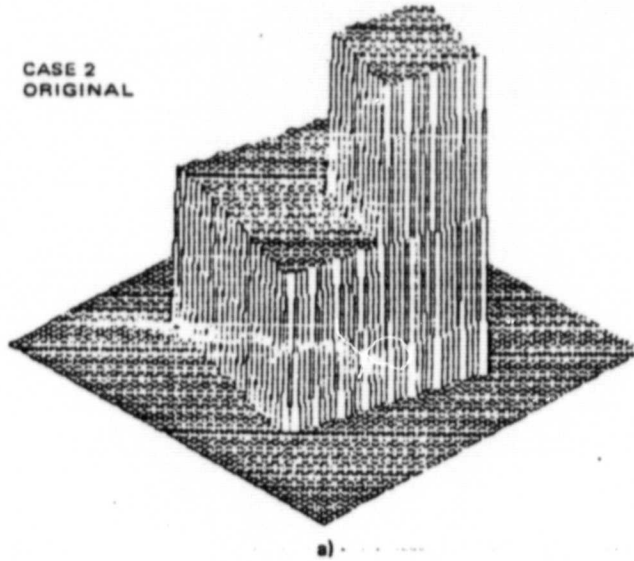
CASE 2
IMAGINARY PART OF
FOURIER TRANSFORM

b)
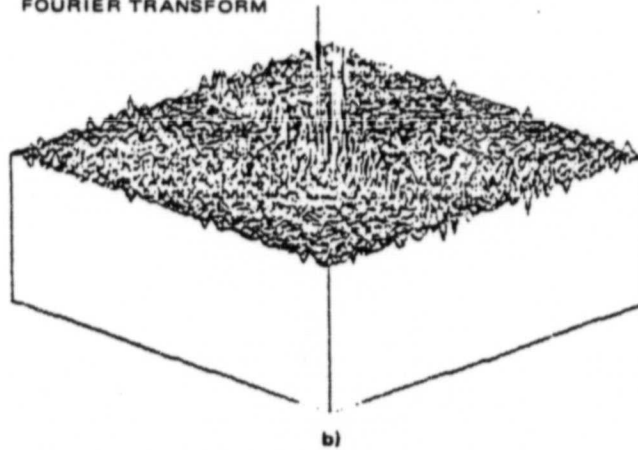
CASE 2
REAL PART OF
FOURIER TRANSFORM

c)
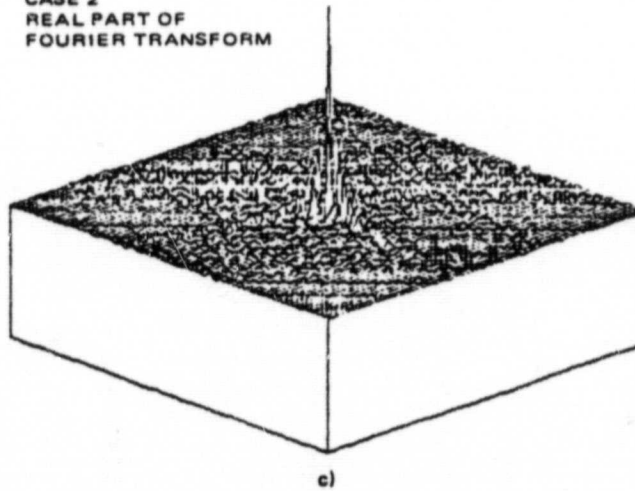
FIGURE 1-5. STACKED-CUBE TEST CASE 2

2. DIRECT FOURIER TRANSFORM TECHNIQUES

2.1 DIRECT FFT INVERSE

Given a set of frequency domain measurements $m_i = F(u_i, v_i)$, $i = 1,2,3, \ldots M$, the simplest reconstruction would be to perform inverse Fourier transform on these measurements. This is to assume that in an under-sampled system, the missing data take on values of zero;

$$\hat{f}(x,y) = \sum_{i=1}^{M} a_i \{m_i \exp[j2\pi(xu_i + yv_i)]\} \qquad (II-2-1)$$

where a_i are chosen similar to the "lag window" in classical power spectral density estimation or the "weighting" on each data m_i. $\hat{f}(x,y)$ so defined is always real since it is known a priori that $F(u,v)$ is conjugated symmetrically. Any set of measurements can be augmented with their complex conjugates at proper negative spatial frequency locations. Obviously, $\hat{f}(x,y)$ obtained from Eq. (II-2-1) will have any kind of resolution only if the measurement coverage of the (u,v) plane is almost complete. Equation (II-2-1) is also lacking in taking into consideration that an admissible reconstruction $\hat{f}(x,y)$ must be nonnegative.

Figures 2-1 and 2-2 show the reconstructions of test cases 1 and 2 for each of the three sampling functions S1, S2, and S3. Sampling functions S2 and S3 represent spatial "brickwall" filters with impulse response of sin(x)/x nature. The spatial convolution of the original cubes with the sin(x)/x response is apparent. Note that although sampling function S3 has approximately the same number of measurements as S1, it enables a much better direct inverse transform reconstruction both subjectively and in squared-error sense. Since both of the test cases have their frequency contents concentrated around zero, S3 perserves much more useful information than S1; and in some sense, is a better data compression algorithm.

2.2 INTERPOLATION OF MEASUREMENTS

There exist two major problems in direct transform techniques. The obvious one is that missing data in an undersampled system are assumed to have zero values. Secondarily, the measurements may not fall on rectangular lattice

CASE 1
SAMPLE FUNCTION S1
DIRECT FFT INVERSE
ERROR = 25.5
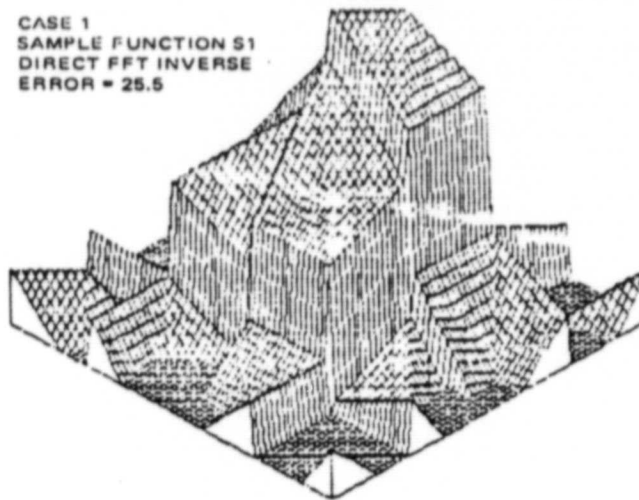
a)

CASE 1
SAMPLE FUNCTION S2
DIRECT FFT INVERSE
ERROR = 12.3
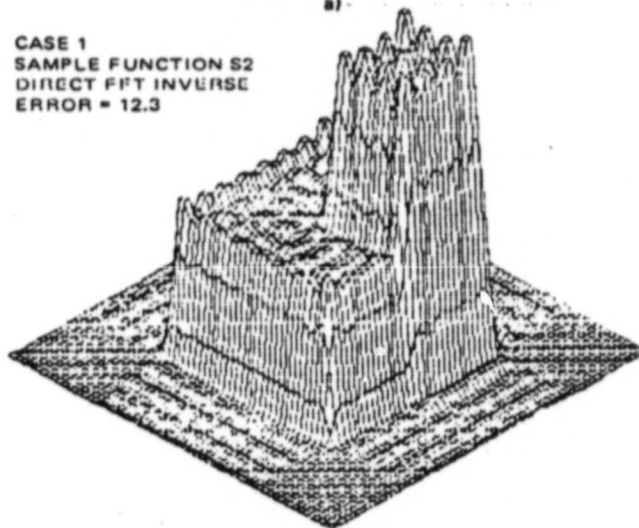
b)

CASE 1
SAMPLE FUNCTION S3
DIRECT FFT INVERSE
ERROR = 19.2
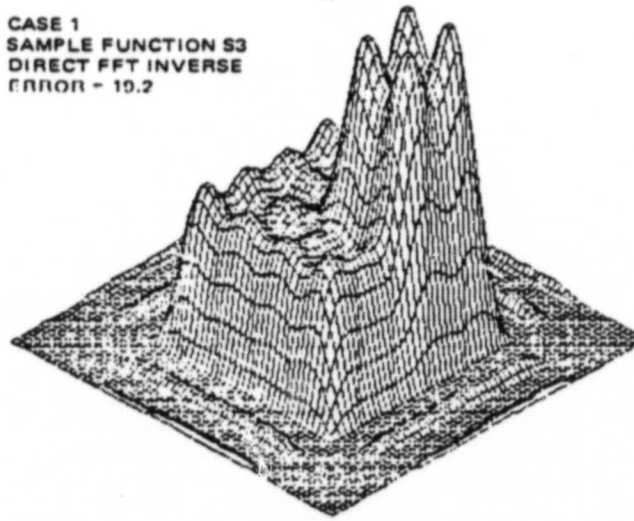
c)

FIGURE 2-1. DIRECT INVERSE TRANSFORM, TEST
CASE 1

CASE 2
SAMPLE FUNCTION S1
DIRECT FFT INVERSE
ERROR = 34.4
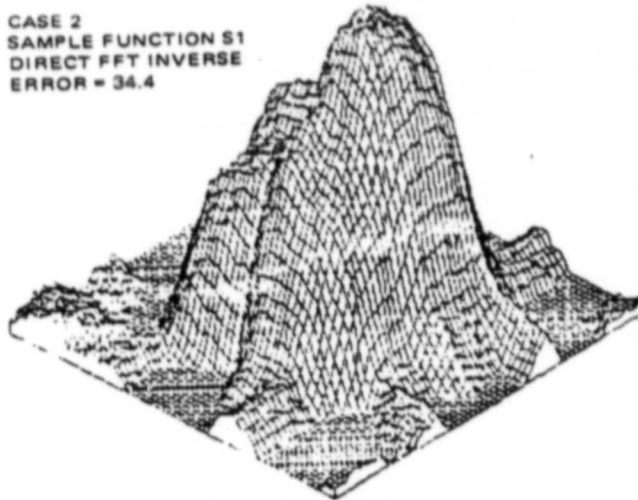
a)

CASE 2
SAMPLE FUNCTION S2
DIRECT FFT INVERSE
ERROR = 17.9
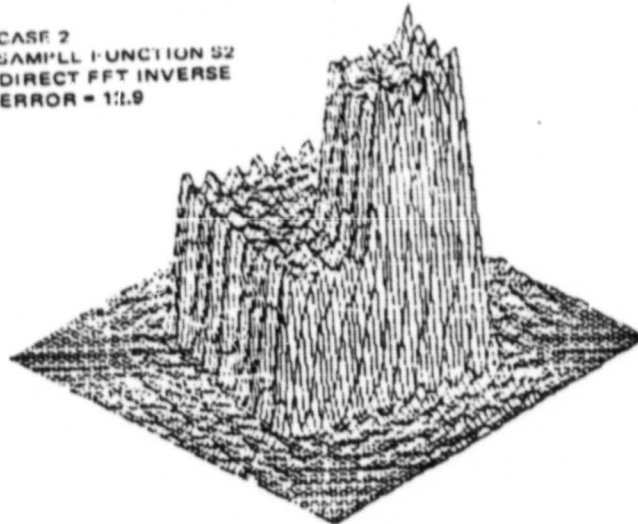
b)

CASE 2
SAMPLE FUNCTION S3
DIRECT FFT INVERSE
ERROR = 19.3
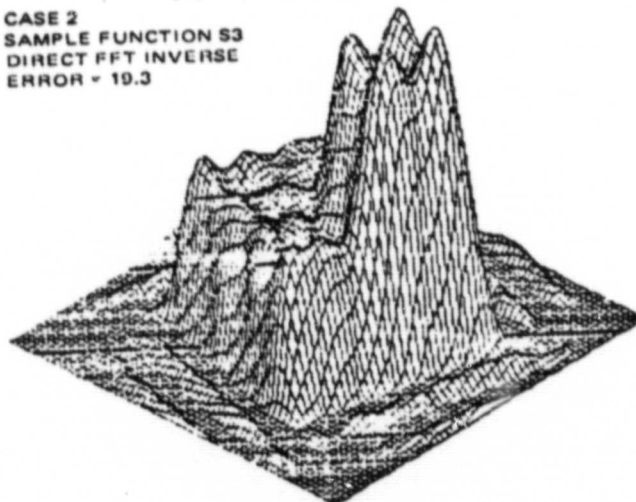
c)

FIGURE 2-2. DIRECT INVERSE TRANSFORM, TEST
CASE 2

TT-2-3

to facilitate computational efficient Fast Fourier-transform (FFT) technique. For example, in rotation synthesis interferometry, measurements lie on a polar rather than rectangular grid. Thompson and Bracewell (Reference 12) studied the method of partitioning the (u,v) plane and performing "cell summing" on the measurements in each cell to obtain a set of data on rectangular lattice. Stark (Reference 13) proved a sampling theorem in polar coordinate where image can be reconstructed from spatial samples. It requires nonuniformly spaced samples correspond to scaled zeros of Bessel functions along radial direction. The problem of transforming these nonuniformly spaced samples onto rectangular grid still exist.

Very little effort has been placed on searching for an optimum interpolation strategy since the multi-λ system is significantly undersampled. Any reconstruction method involving severe interpolation introduces intolerable artifacts. A simple method is tested (Reference 14).

Assuming that the frequency data along axes and diagonals are sufficiently sampled (beyond the Nyquist rate), therefore the entire axes and diagonals are completely specified via Nyquist sampling theorem. Measurements can be arranged in (θ,f) plane as illustrated by Figure 2-3. Every point on the rectangular grid inside the circle with radius f_{max} can be interpolated via a simple bicubic spline on (θ,f) (Reference 15). Computational efficient FFT algorithms then can be applied on interpolated data. Since frequency measurements are complex, the real and the imaginary parts are interpolated independently. Figure 2-4 shows the resulting reconstruction on Case 1 and sampling function S1. Figure 2-5 has a sampling function such that angular θ is specified every 2 degrees from -180 to 180 as opposed to every 45 degrees for S1. The two figures clearly indicate the relationship between artifacts and severity of interpolations. It is interesting to note that in order to completely specify every point on the 64x64 rectangular grid without any interpolation, each frequency quadrant (i.e., 0-90 degree) must contain 651 distinct nonuniformly spaced angles. It roughly corresponds to an angular resolution of 0.14 degrees.

a) ORIGINAL RECTANGULAR LATTICE

b) (θ, F) TRANSFORMATION FOR BICUBIC SPLINE
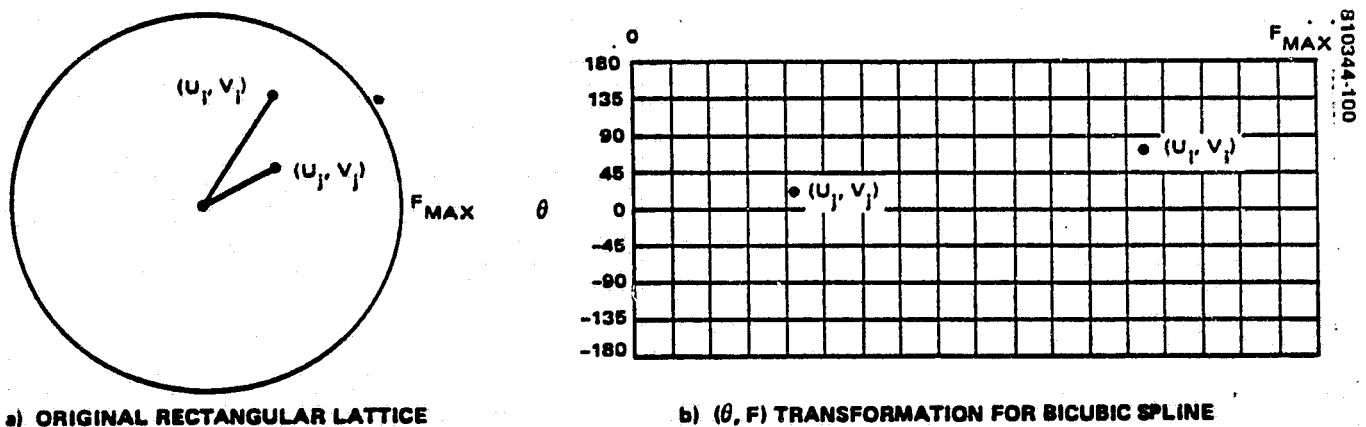
FIGURE 2-3. BICUBIC SPLINE INTERPOLATION
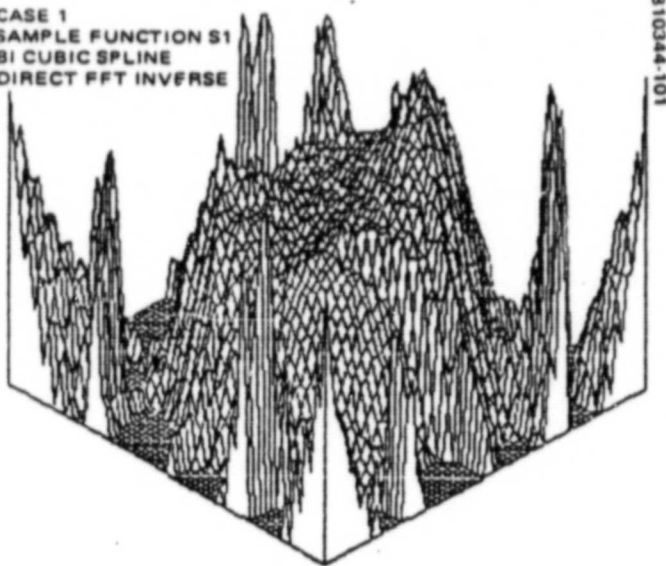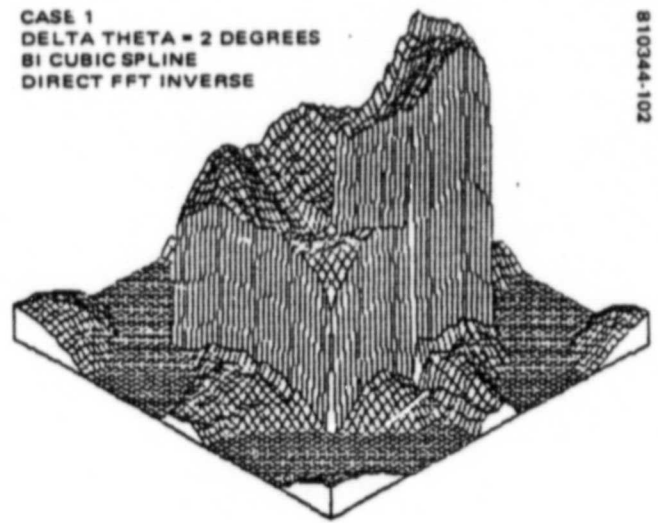
FIGURE 2-4. INTERPOLATION ARTIFACTS, $\theta = 45^{\circ}$

FIGURE 2-5. INTERPOLATION ARTIFACTS, $\theta = 2^{\circ}$

2.3 INVERSE FILTERING AND WIENER FILTERING

Inverse and Wiener filtering techniques are two classical restoration methods. Their usefulness is mainly to remove linear PSF or MTF distortion. Wiener filter is, superior to inverse filter when the data are noisy and it reduces down to inverse filter in the noiseless case. Extending these classical filtering techniques from one dimension to two dimensions involves applying theory of Toeplitz and circulant matrices (Reference 1). Results quoted here are quite intuitive, however.

Let $F(u,v)$ be the complex two-dimensional Fourier transform of $f(x,y)$ and let $H(u,v)$ be the system transfer function. Let $G(u,v)$ be the measurements field and

$$G(u,v) = H(u,v)F(u,v) \qquad (II-2-2)$$

If $H(u,v)$ is nonsingular everywhere over the (u,v) plane, then $F(u,v)$ can be estimated from $G(u,v)$ by

$$\hat{F}(u,v) = G(u,v)/H(u,v)$$

and

$$\hat{f}(x,y) = \iint \hat{F}(u,v) \exp[j2\pi(xu+yv)]\,du\,dv \qquad (II-2-3)$$

If all measurement (u,v) pairs on G(u,v) are on rectangular lattices, $\hat{f}(x,y)$ can be computed by inverse FFT algorithm by discretizing Eq. (II-2-3). This inverse filtering technique is also called least-square method.

Apparently, if $H(u,v)$ = S1 as defined in Section 1, H(u,v) is singular and takes on zero values over most of the (u,v) plane. Although G(u,v) = 0 also whenever H(u,v) = 0, G(u,v)/H(u,v) gives an indetermined ratio. To show an example where inverse filtering works for nonsingular H(u,v), let S1'.be such that

$$S1'\{F(u,v)\} = \begin{cases} F(u,v) & u=v, u=-v, u=0, v=0 \\ \\ \epsilon \neq 0 & \text{else, where } \epsilon \text{ small} \end{cases} \tag{II-2-4}$$

or,

$$H(u,v) = \begin{cases} 1 & u=v, u=-v, u=0, v=0 \\ \\ \epsilon & \text{else} \end{cases} \tag{II-2-5}$$

Obviously Eq. (II-2-2) and (II-2-3) yield $\hat{F}(u,v)$ = F(u,v). The inverse filter 1/H(u,v) restores f(x,y) perfectly.

There are two inherent shortcomings in this technique. There will be artifacts around the edge of the reconstructed image from the circular convolution wrapped around effect in calculating Eq. (II-2-3). Also if the data are noisy, the noise components along with measurements are amplified by 1/H(u,v) when H(u,v) is small. Figure 2-6(b) shows the noisy reconstruction using inverse filter 1/H(u,v) where H(u,v) is defined by Equation (II-2-5).

Wiener filter is derived to minimize mean-square-error between f and $\hat{f}$ assuming the knowledge of both the image and the noise statistics. It can be shown that (Reference 1)

$$\hat{F}(u,v) = \frac{H^*(u,v)G(u,v)}{|H(u,v)|^2 + [P_n(u,v)/P_f(u,v)]}$$

and

$$\hat{f}(x,y) = \iint F(u,v) \exp[j2\pi(xu+yv)] \, du \, dv \tag{II-2-6}$$

where $P_n(u,v)$ and $P_f(u,v)$ are the power spectral densities of the noise and the image respectively; and H^* denotes complex conjugate.

810344-103
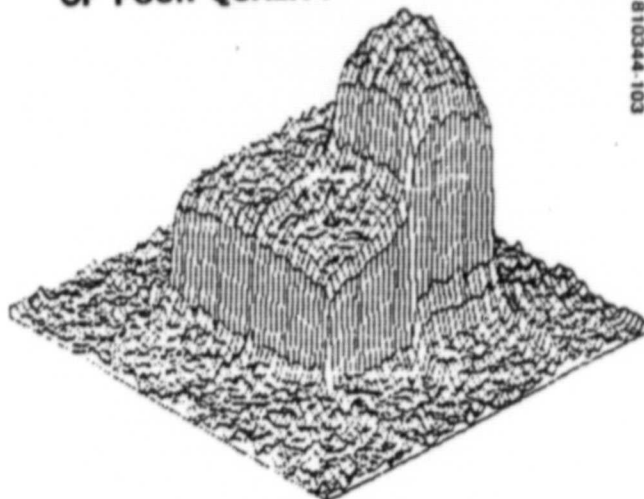
a) WIENER FILTER RECONSTRUCTION
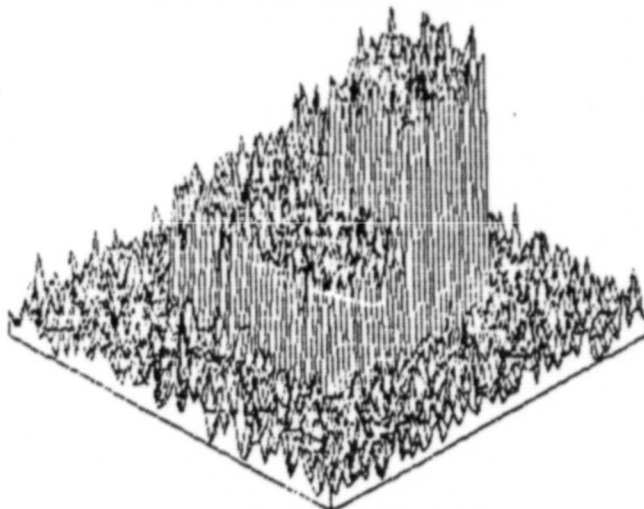
b) INVERSE FILTER RECONSTRUCTION
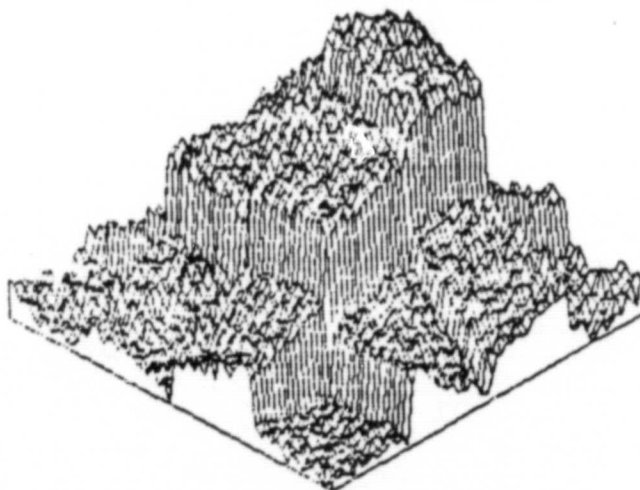
c) DIRECT FFT^{-1} RECONSTRUCTION

FIGURE 2-6. EXAMPLE OF INVERSE AND WIENER FILTERING

There is no ill-conditioned behavior of singularity problem associated with Wiener filter. Note that if $H(u,v) \to 0$, $\hat{F}(u,v) \to 0$ also. This implies that missing data are treated as zeros as in the case of direct inverse transform. If $P_n \to 0$, i.e., noiseless, Wiener filtering of Eq. (II-2-6) reduces to Eq. (II-2-3), the inverse filtering approach. Furthermore, if $P_f \to 0$ (no signal) $F(u,v) \to 0$. This implies that one can not expect to recover information at (u,v) where noise is dominant.

Figure 2-6(a) illustrates the improvement on noise problem using Wiener filtering over inverse filtering technique. P_n is assumed to be white Gaussian noise and P_f is estimated from the two-dimensional FFT of the origianl test pattern. It is well known that Wiener filter is a smoothing filter; reconstructed image appears flat and fine details are blurred.

3. MAXIMUM ENTROPY RECONSTRUCTION

3.1 DISCUSSION OF MAXIMUM ENTROPY METHOD

In view of inadequate reconstruction techniques when the problem is
singular, Wernecke proposed and implemented an iterative algorithm utilizing
the principle of maximum entropy [References 8 through 11]. When the recon-
struction problem is singular, one has to find a solution of image that is
consistent with the known measurements. Of course, solution neither has to
exist nor has to be unique. A measure to pick the "best" solution is needed.
In addition, the a priori knowledge that the solution (image) being non-
negative is a very strong constraint and should be considered in the recon-
struction process.

Maximum entropy principle has been applied successfully to estimate the
power spectral density $S(v)$ of a stationary random process from measurements
of the correlation function. There is no direct mapping between $S(v)$ and
the entropy rate of the random process in the Shannon sense. However, the
random process with spectral density $S(v)$ can be viewed as the output process
of a linear filter with squared magnitude frequency response of $S(v)$ and
driven by a white input process. Bartlett showed that the entropy gain of
the filter (difference of the input and output entropies) can be expressed
by [References 16 and 17]

$$E = \int \log S(v) \, dv \qquad \text{(II-3-1)}$$

It is not required to know the complete statistics of the input except that
it is a white process. The entropy of the output process is maximized if
the entropy gain of the filter E is made as large as possible. Naturally it
is necessary to have the constraint that the estimated spectrum $\hat{S}(v)$ has to
be consistent to the measured data. Maximum entropy method can be viewed as
the "most random" extension to the known data. In the case of one-dimensional
MEM spectral estimation, the solution is achieved via polynomial spectral
factorization. Part I discusses few pros and cons of this method. Two-
dimensional extension only has limited success [Reference 4].

Wernecke observed that there are great similarities between reconstructing radio brightness distribution in Fourier synthesis and spectral estimations. Both reconstructions are subject to nonnegative constraint and are related to available measurement data via Fourier-transform relationship. Defining a similar entropy measure H,

$$H = \iint_D \log f(x, y) \, dx \, dy \qquad \text{(II-3-2)}$$

the problem can be formulated as

$$\max \iint_D \log \hat{f}(x, y) \, dx \, dy$$

such that

$$\iint_D \hat{f}(x, y) \exp[-j2\pi(xu_i + yv_i)] \, dx \, dy = m_i(u_i, v_i)$$

and

$$\hat{f}(x, y) \geq 0 \quad \text{for all } (x, y) \text{ in } D \qquad \text{(II-3-3)}$$

where D is the field of view. If the measurements are not noiseless nor error free, m_i, $i = 1, 2, \ldots M$ become $\hat{m}_i$, and additional constraint

$$\sum_{i=1}^{M} \frac{1}{\sigma_i^2} |m_i - \hat{m}_i|^2 = M \qquad \text{(II-3-4)}$$

has to be added to Eq. (II-3-3). σ_i^2's represent the variances of the measurements. This constrained optimization problem may not have an explicit solution and is hard to solve numerically. Furthermore, given a set of measurements, there may not be a nonnegative solution that optimizes Eq. (II-3-3). The problem can be turned into an unconstrained optimization of

$$\max \iint_D \log \hat{f}(x, y)\ dx\ dy - \lambda \sum_{i=1}^{M} \frac{1}{\sigma_i^2} |m_i - \hat{m}_i|^2 \qquad (\text{II-3-5})$$

where λ plays the role of Lagrange multiplier and is chosen so Eq. (II-3-4) is satisfied to a sufficient accuracy. Equation (II-3-5) has a nonnegative solution $\hat{f}(x, y)$ for any $\lambda > 0$. A numerical algorithm can be developed by partitioning $f(x, y)$ into pixels (discretize the integrals) so that Eq. (II-3-5) can be solved iteratively [Reference 8]. A user's manual and program listing supplied by Wernecke is attached in the Volume II, Appendix.

3.2 EXAMPLES

Figures 3-1 through 3-3 show the reconstructions of the two stacked cube test cases with sampling functions S1, S2, and S3. Since all three sampling functions have large number of zero values, the entropy given by Eq. (II-2-3) is modified slightly to stabilize the computations:

$$H' = \iint \log (1 + f(x, y))\ dx\ dy \qquad (\text{II-3-6})$$

Approximations of Shannon entropy associated with these images and reconstructions are given along with the squared-errors. They are computed by using relative frequency approximation of the probability densities where

$$P_i = f_i/F \qquad f_i = \text{ith pixel intensity}$$

$$F = \sum_i f_i \text{ is the total intensity}$$

CASE 1
SAMPLE FUNCTION S1
MAXIMUM ENTROPY
4 ITERATIONS
ERROR = 22.5
H_1 = 2.3
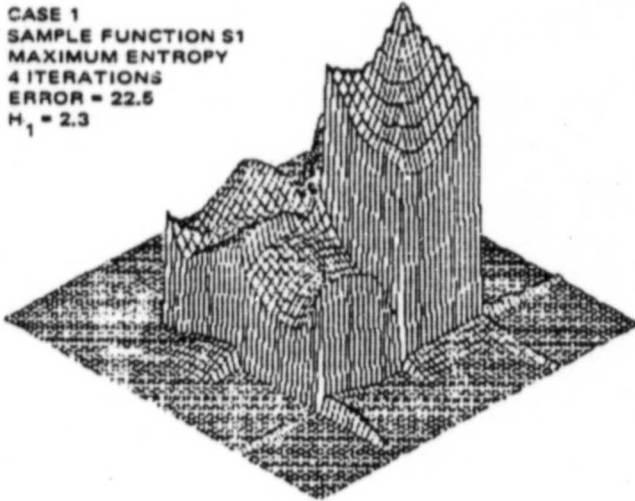
CASE 1
SAMPLE FUNCTION S1
MAXIMUM ENTROPY
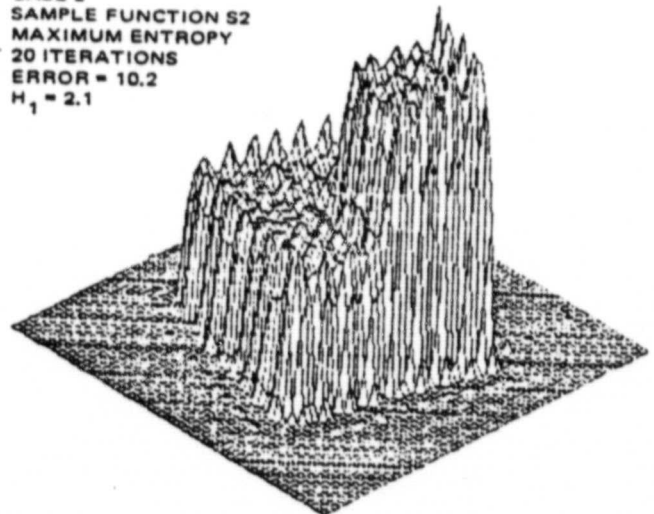20 ITERATIONS
ERROR = 7.4
H_1 = 1.8

810344-104
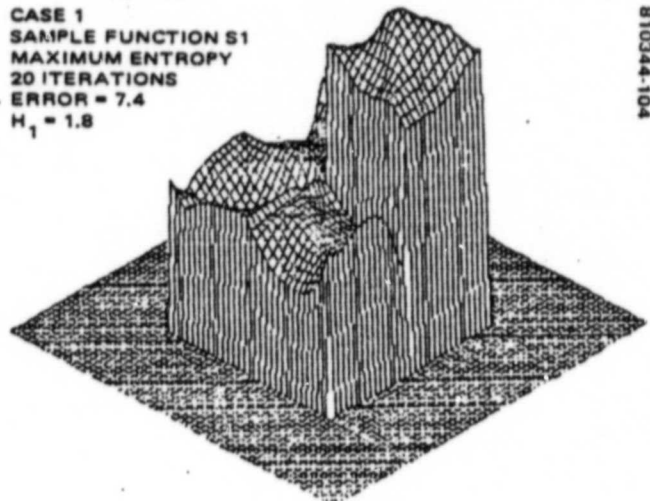
a) TEST CASE 1 ORIGINAL H_1 = 0.7

CASE 2
SAMPLE FUNCTION S1
MAXIMUM ENTROPY
4 ITERATIONS
ERROR = 37.4
H_1 = 2.7
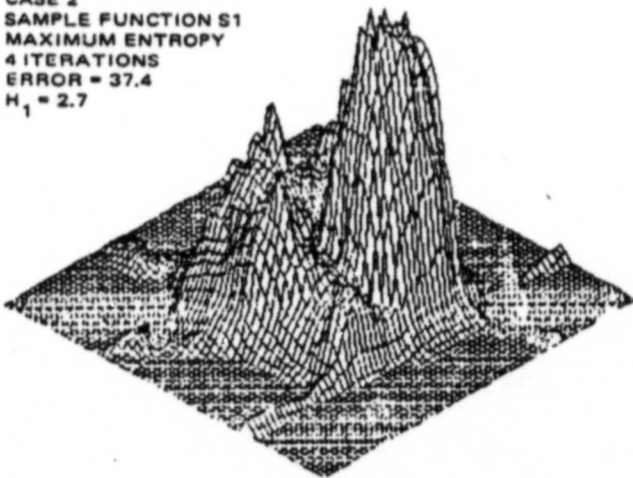
CASE 2
SAMPLE FUNCTION S1
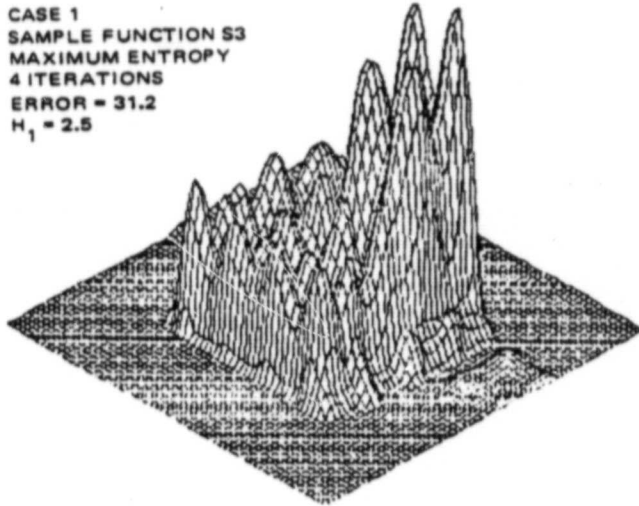MAXIMUM ENTROPY
20 ITERATIONS
ERROR = 30.3
H_1 = 2.7
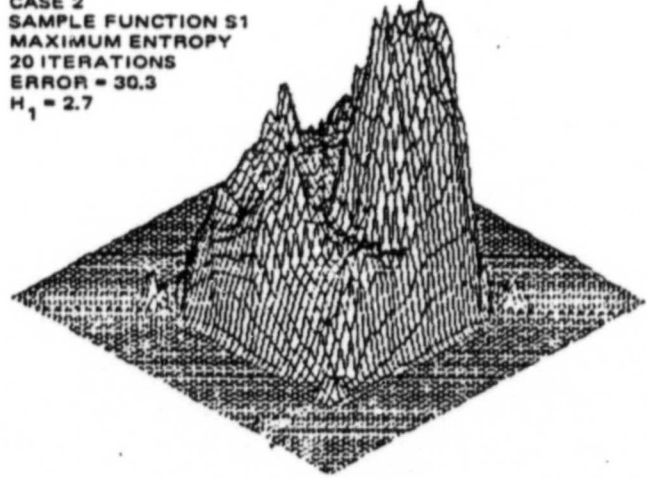
b) TEST CASE 2 ORIGINAL H_1 = 0.7
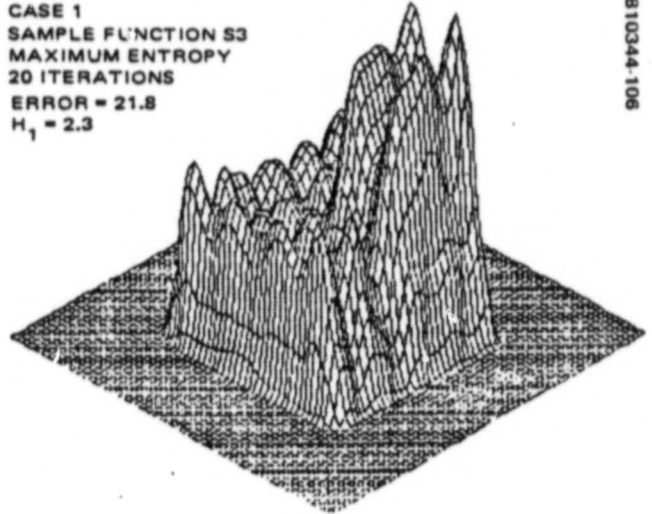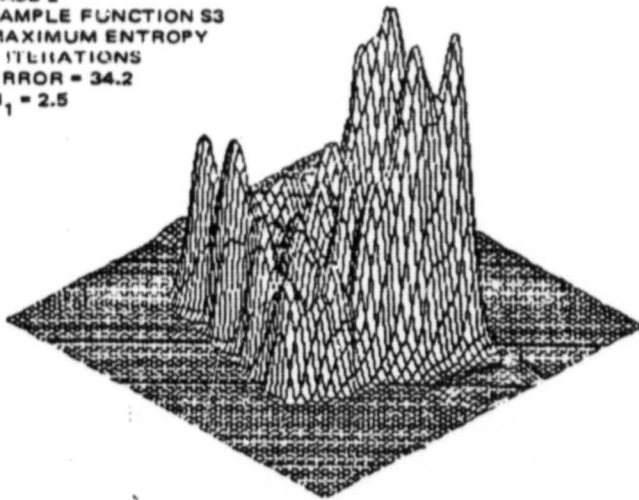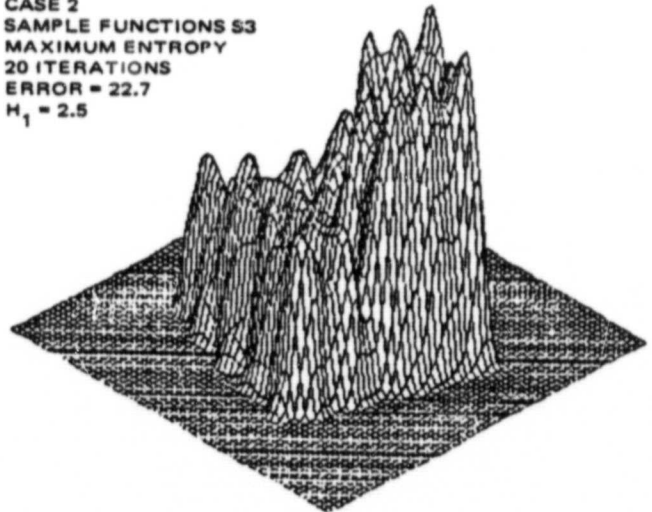
FIGURE 3-1. MEM RECONSTRUCTION, SAMPLING FUNCTION S1

CASE 1
SAMPLE FUNCTION S2
MAXIMUM ENTROPY
4 ITERATIONS
ERROR = 13.8
H_1 = 2.3
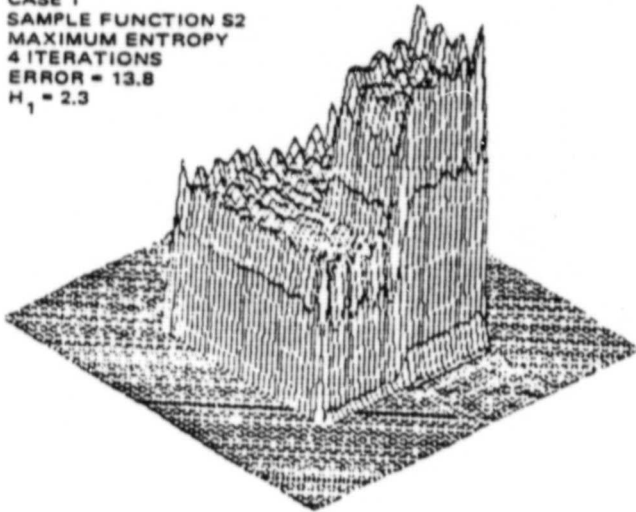
CASE 1
SAMPLE FUNCTION S2
MAXIMUM ENTROPY
20 ITERATIONS
ERROR = 8.4
H_1 = 1.9

810344-105
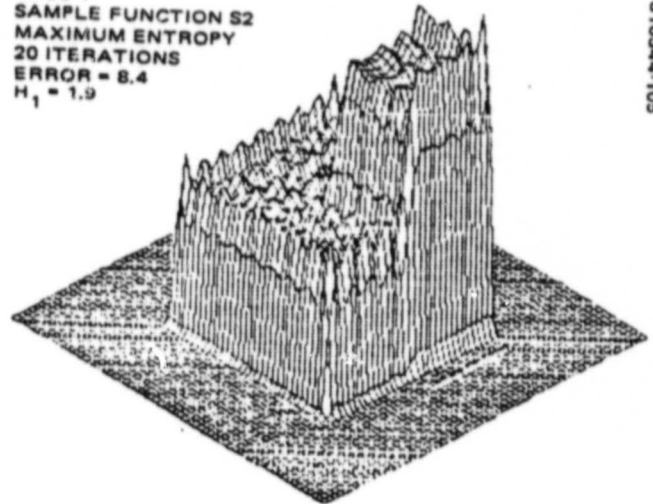
a) TEST CASE 1 ORIGINAL H_1 = 0.7

CASE 2
SAMPLE FUNCTION S2
MAXIMUM ENTROPY
4 ITERATIONS
ERROR = 15.4
H_1 = 2.4
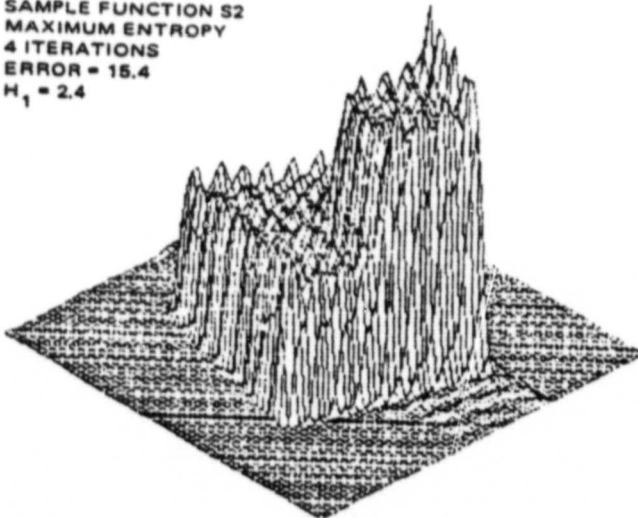
CASE 2
SAMPLE FUNCTION S2
MAXIMUM ENTROPY
20 ITERATIONS
ERROR = 10.2
H_1 = 2.1

b) TEST CASE 2 ORIGINAL H_1 = 0.7

FIGURE 3-2. MEM RECONSTRUCTION, SAMPLING FUNCTION S2

a) TEST CASE 1 ORIGINAL $H_1 = 0.7$

b) TEST CASE 2 ORIGINAL $H_1 = 0.7$

FIGURE 3-3. MEM RECONSTRUCTION, SAMPLING FUNCTION S3

and

$$H_1 = -\sum_i p_i \log(p_i)$$

(II-3-7)

In all the cases, MEM method yields stable solutions that are both good in the subjective and squared-error sense. When the data are not severely undersampled (S2), it converges rather quickly. Table 3-1 summarizes squared-errors for MEM and direct-transform reconstructions on the two test cases. It is interesting to note that for Case 2 and S1 (Figure 3-1), the phase rotation of the original is lost in the reconstruction.

TABLE 3-1.   SUMMARY OF SQUARED-ERROR FOR
             TEST CASE 1 AND 2

		S1	S2	S3
CASE 1	MEM	7.4	8.4	21.8
	FFT^{-1}	25.5	12.3	19.2
CASE 2	MEM	30.3	10.2	22.7
	FFT^{-1}	34.4	12.9	19.3

3.3   ALTERNATIVE SOLUTIONS TO MAXIMUM ENTROPY METHOD

Wernecke also obtained a set parametric expression using Calculus of Variations for various entropy criteria [Reference 8]. Tables 3-2 and 3-3 summarize his results for both error-free measurements and measurements with errors. Reconstruction can also be done based on numerical solution of these nonlinear parametric equations. Iterative method given in Section 3.2 and Appendix L is more preferable since the image, $f(x, y)$, is manipulated directly rather than the parameters which has no bearing on the image quality. Any one of these parametric equations can be incorporated into the MEM algorithm with only minor programming modifications.

TABLE 3-2. PARAMETRIC EXPRESSIONS FOR ERROR-INTOLERANT MAXIMUM
ENTROPY IMAGE RECONSTRUCTION

ENTROPY/SMOOTHNESS CRITERION	PARAMETERIZED RECONSTRUCTION
$$\iint_D \ln f(x, y)\, dx\, dy$$	$$\frac{1}{\displaystyle\sum_{i=1}^{M} \lambda_i h_i(x, y)}$$
$$-\iint_D [f(x, y)/F]\, \ln[f(x, y)/F]\, dx\, dy \quad \text{F known}$$	$$\exp\left[-1 - F\sum_{i=1}^{M} \lambda_i h_i(x, y)\right]$$
$$-\iint_D [f(x, y)/F]\, \ln[f(x, y)/F]\, dx\, dy \quad \text{F unknown}$$	$$\exp\left[-H_0/F - F\sum_{i=1}^{M} \lambda_i h_i(x, y)\right]$$
$$-\iint_D f(x, y)\, \ln f(x, y)\, dx\, dy$$	$$\exp\left[-1 - \sum_{i=1}^{M} \lambda_i(x, y)\right]$$
$$\iint_D \sqrt{f}(x, y)\, dx\, dy$$	$$\frac{1}{\left[2\displaystyle\sum_{i=1}^{M} \lambda_i h_i(x, y)\right]^2}$$

PARAMETER DEFINITION

$$m_k = \iint_D h_k(x, y)\, \hat{f}(x, y)\, dx\, dy \qquad k = 1, 2, \ldots, M$$

$$H_0 = -\iint_D \hat{f}(x, y)\, \ln \hat{f}(x, y)\, dx\, dy$$

$$F = \iint_D \hat{f}(x, y)\, dx\, dy$$

TABLE 3-3. PARAMETRIC EXPRESSIONS FOR ERROR-TOLERANT MAXIMUM ENTROPY
IMAGE RECONSTRUCTIONS

ENTROPY/SMOOTHNESS CRITERION	PARAMETERIZED RECONSTRUCTION
$\iint_D \ln f(x, y)\, dx\, dy$	$\dfrac{1}{2\lambda \sum\limits_{i=1}^{M} \dfrac{\Delta_i}{\sigma_i^2} h_i(x, y)}$
$-\iint_D [f(x, y)/F]\, \ln[f(x, y)/F]\, dx\, dy$	$\exp\left[-H_0/F - 2\lambda F \sum\limits_{i=1}^{M} \dfrac{\Delta_i}{\sigma_i^2} h_i(x, y)\right]$
$-\iint_D f(x, y)\, \ln f(x, y)\, dx\, dy$	$\exp\left[-1 - 2\lambda \sum\limits_{i=1}^{M} \dfrac{\Delta_i}{\sigma_i^2} h_i(x, y)\right]$
$\iint_D \sqrt{f(x, y)}\, dx\, dy$	$\dfrac{1}{\left[4\lambda \sum\limits_{i=1}^{M} \dfrac{\Delta_i}{\sigma_i^2} h_i(x, y)\right]^2}$

PARAMETER DEFINITION

$$\Delta_k = \iint_D h_k(x, y)\, \hat{f}(x, y)\, dx\, dy - m_k \qquad k = 1, 2, \ldots, M$$

$$H_0 = -\iint_D \hat{f}(x, y)\, \ln \hat{f}(x, y)\, dx\, dy$$

$$F = \iint_D \hat{f}(x, y)\, dx\, dy$$

$$M = \sum_{i=1}^{M} \frac{\Delta_i^2}{\sigma_1^2} \qquad \text{(if the value of } \lambda \text{ is not chosen \underline{a priori}} \atop \text{by the user)}$$

4. RADIOMETER IMAGE TEST CASES

Three radiometer images are chosen to test the reconstruction algorithms and sampling functions. Test Image 1 mainly consists of fields and roads. Test Image 2 has a mixture of fields and city area. Test Image 3 has a more detailed city outline. The original images are shown in Figures 4-4(a), 4-5(a), and 4-6(a). Figures 4-1 through 4-3 are the spatial frequency distributions of the images obtained by two-dimensional FFT. Frequency-domain data are simulated by applying sampling function S1, S2, and S3 to these Fourier-transformed images.

Figures 4-4 through 4-6 show the reconstruction techniques under investigation for sampling function S1, the simulated multi-λ system. Figures 4-7 through 4-9 show the results of sampling function S2 and Figures 4-10 through 4-12 show the effects of sampling functions S3. Table 4-1 summarizes the results in terms of squared-error for all the sampling functions and two reconstructions. The low error values are somewhat deceiving since the original images have relatively low dynamic range. Figures 4-13 and 4-14 rearrange the displays of the results of Test Image 1 to contrast the effects between various sampling strategies.

As in the case of stacked-cube, S3 is clearly a better sampling function than S1 for approximately same number of data. As a matter of fact, under the assumption that the power spectral density of radiometer images is unimodal ($\Phi(p) = p^{-5/2}$), rate distortion theory discussed in Section 1.2 suggests that one should try to preserve as much low frequency region as possible.

Maximum entropy method did not provide any improvements in all the cases tested except for stacked-cube case with S1. Its usefulness should not be discounted, however. The method is iterative and in all cases, it converged rather quickly. Table 4-2 shows the convergent reconstruction of Image 1 with sampling function S1. In Table 4-3, MEM is given all the spatial frequency samples. It converges to a solution with insignificant error. Since the MEM has built in measurement error/noise and nonnegative reconstruction constraints, it may be a more viable restoration technique than direct inverse transforms even the system is not severely undersampled.

TEST IMAGE 1
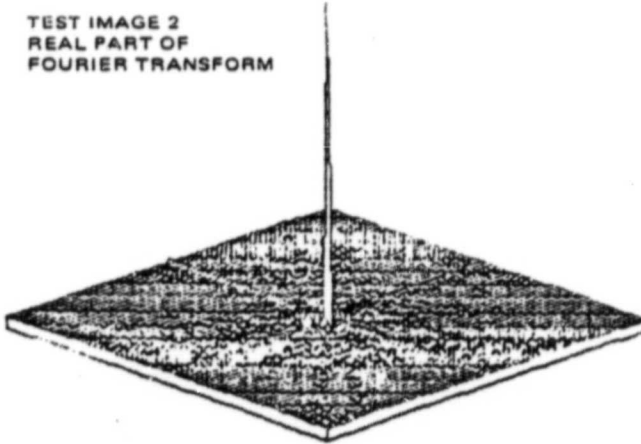REAL PART OF
FOURIER TRANSFORM
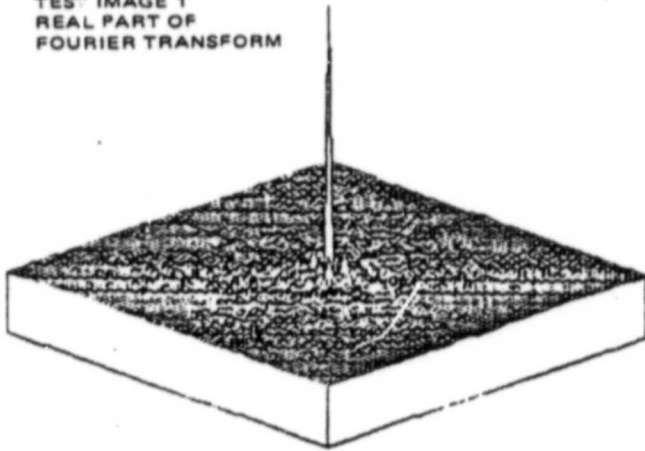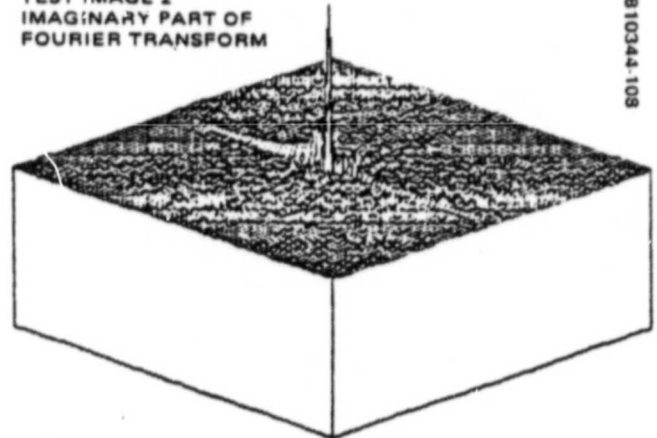
TEST IMAGE 1
IMAGINARY PART OF
FOURIER TRANSFORM
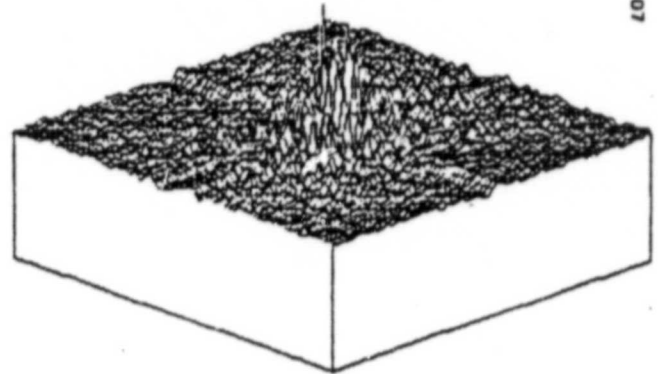
810344-107

FIGURE 4-1. FOURIER TRANSFORM OF IMAGE 1

TEST IMAGE 2
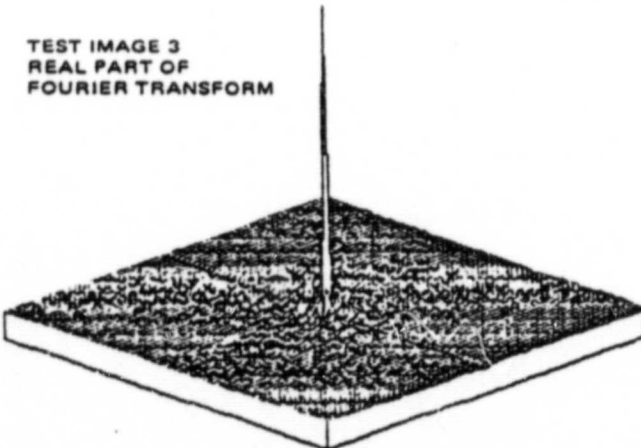REAL PART OF
FOURIER TRANSFORM
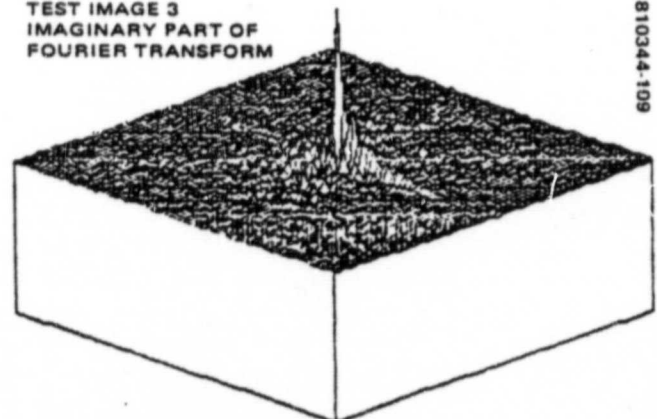
TEST IMAGE 2
IMAGINARY PART OF
FOURIER TRANSFORM

810344-108

FIGURE 4-2. FOURIER TRANSFORM OF IMAGE 2

TEST IMAGE 3
REAL PART OF
FOURIER TRANSFORM
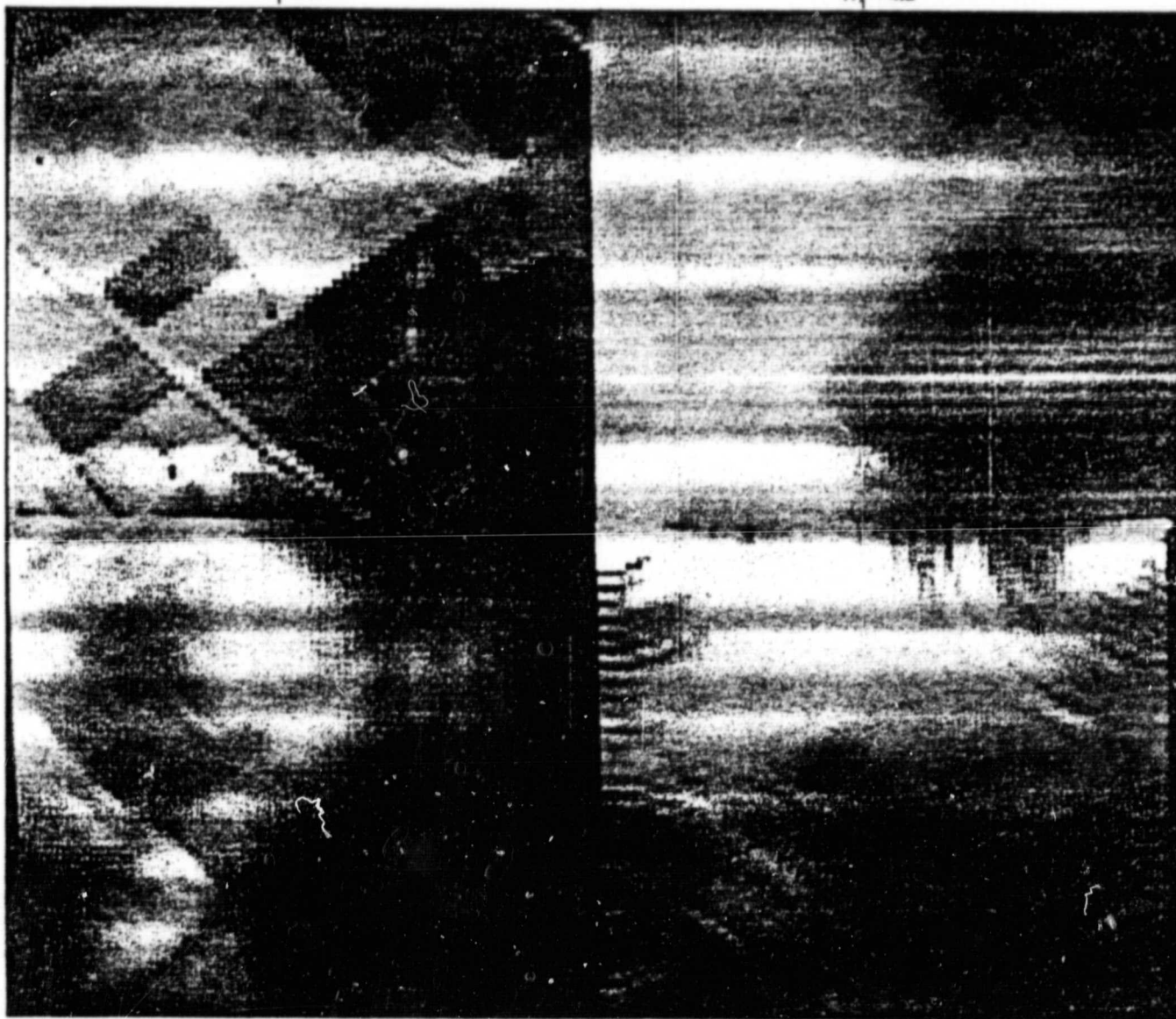
TEST IMAGE 3
IMAGINARY PART OF
FOURIER TRANSFORM

810344-109
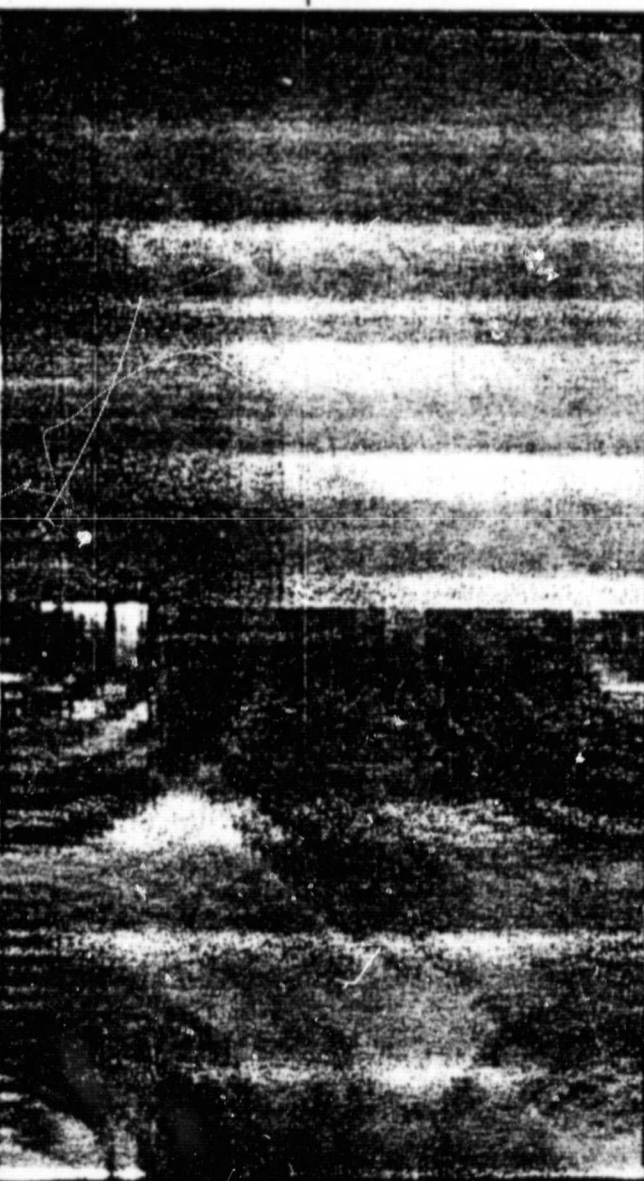
FIGURE 4-3. FOURIER TRANSFORM OF IMAGE 3

a) ORIGINAL
$H_1 = 4.4$
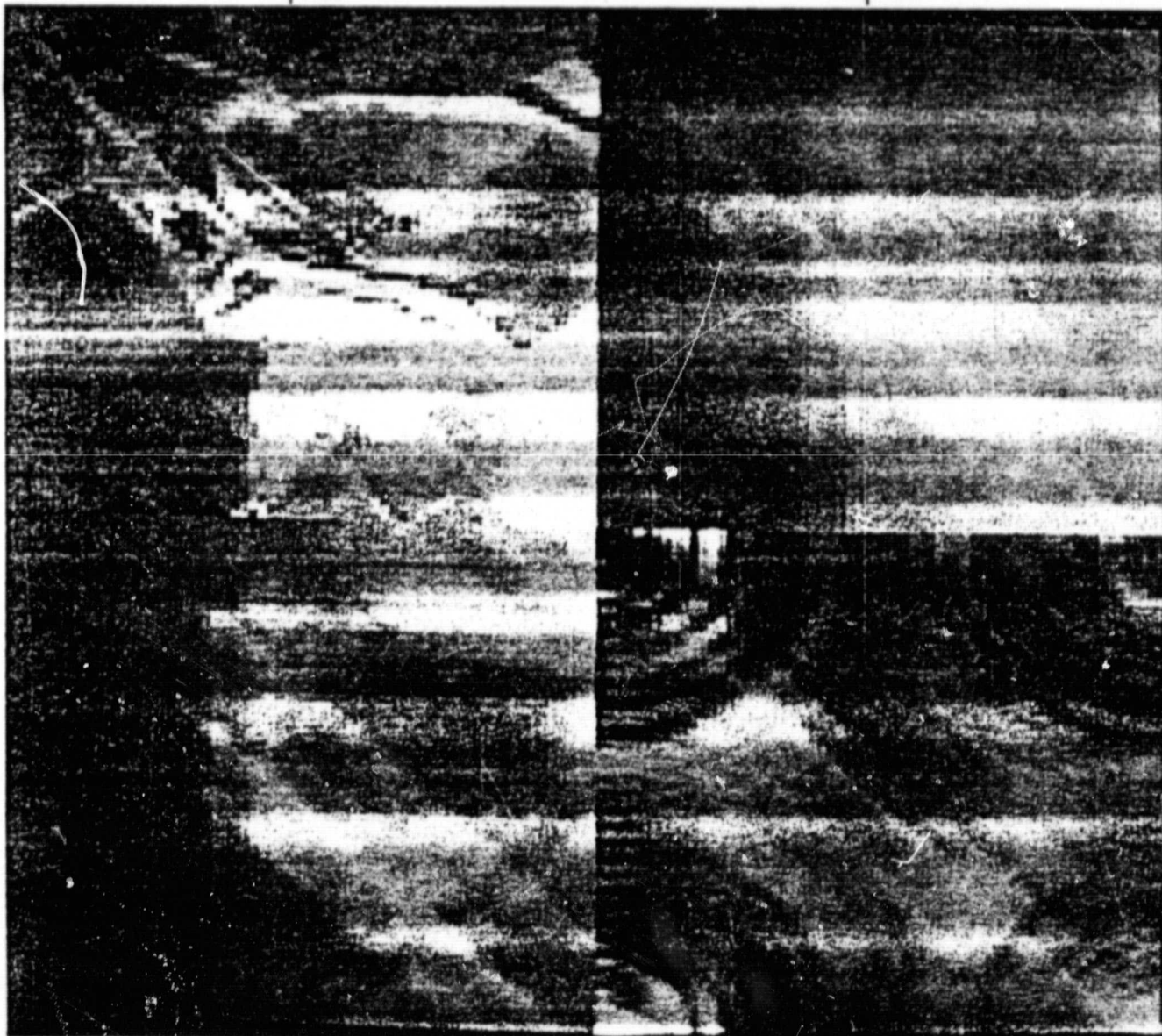
b) INVERSE FFT
ERROR = 14.5
$H_1 = 4.2$

ERROR = 16.7
$H_1 = 4.0$
c) MAXIMUM ENTROPY

ERROR = 23.8
$H_1 = 4.3$
d) BI-CUBIC SPLINE

FIGURE 4-4. SAMPLING FUNCTION S1 APPLIED ON IMAGE 1

a) ORIGINAL

$H_1 = 4.2$

b) INVERSE FFT

ERROR = 9.8
$H_1 = 4.3$
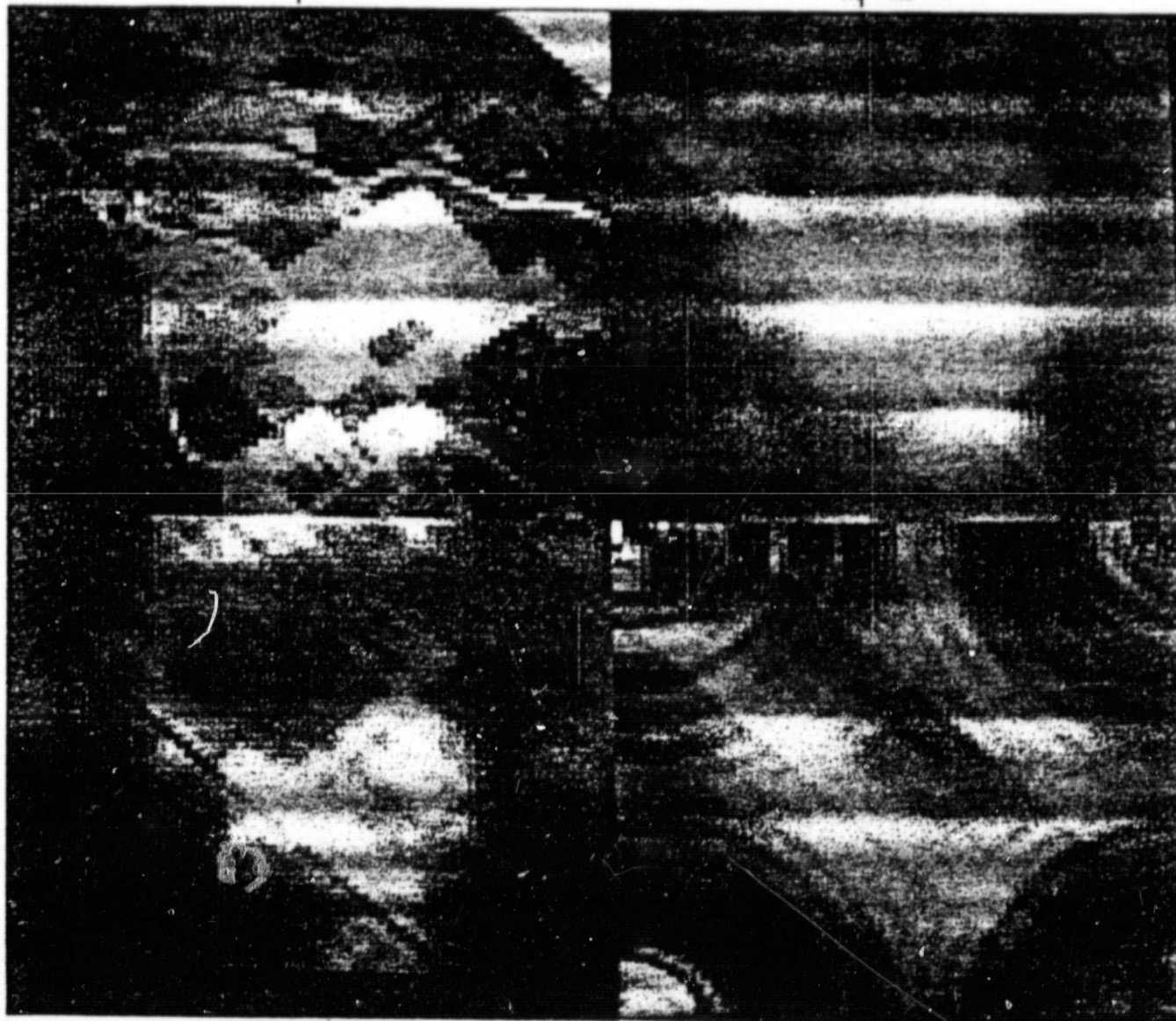
ERROR = 15.2
$H_1 = 4.0$

c) MAXIMUM ENTROPY

ERROR = 26.3
$H_1 = 4.4$

d) BI-CUBIC SPLINE

FIGURE 4-5. SAMPLING FUNCTION S1 APPLIED ON IMAGE 2

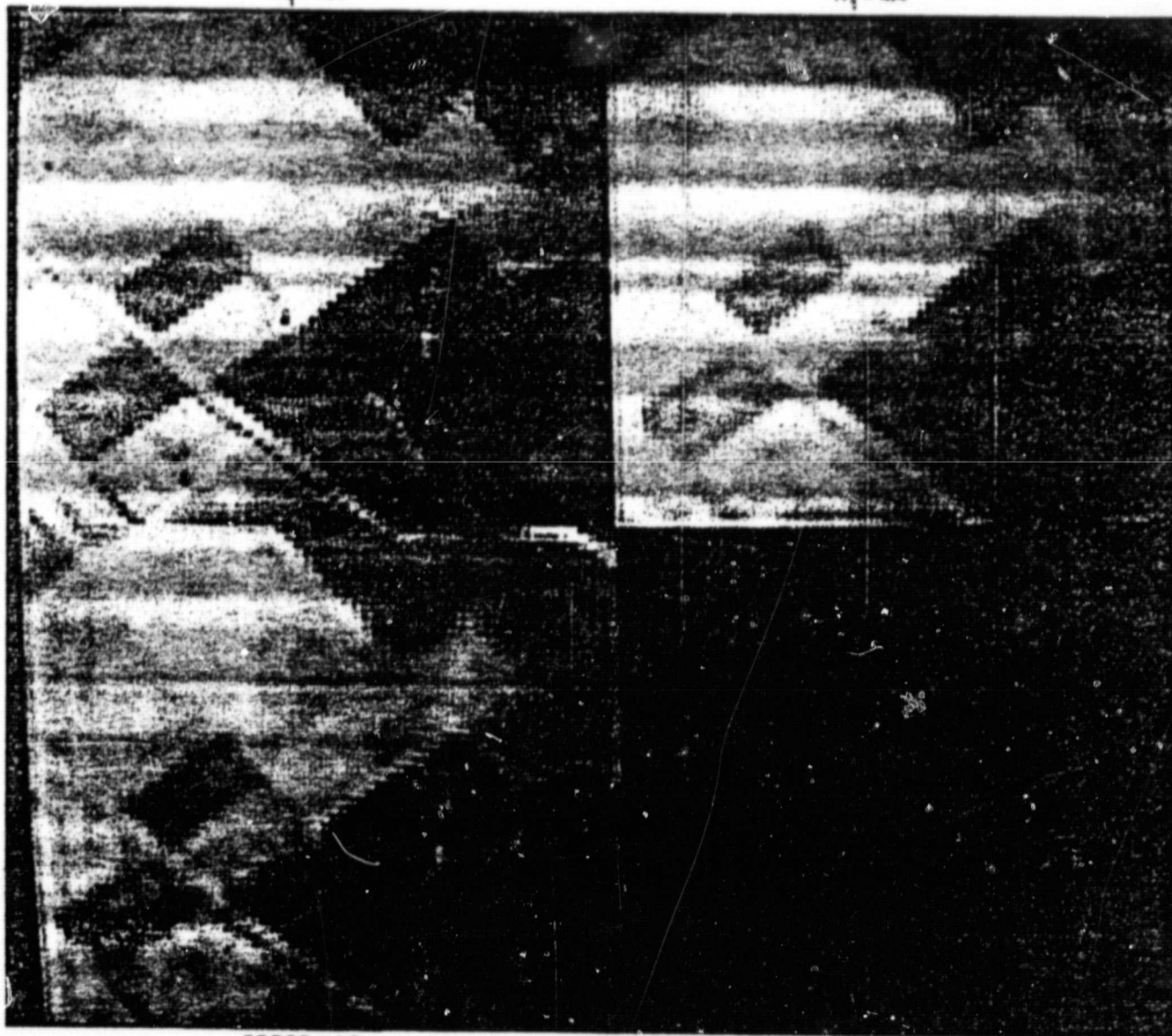FIGURE 4-6. SAMPLING FUNCTION S1 APPLIED ON IMAGE 3

a) ORIGINAL

$H_1 = 4.4$

b) INVERSE FFT

ERROR = 4.8
$H_1 = 4.4$

ERROR = 4.2
$H_1 = 4.2$

c) MAXIMUM ENTROPY

FIGURE 4-7. SAMPLING FUNCTION S2 APPLIED ON IMAGE 1
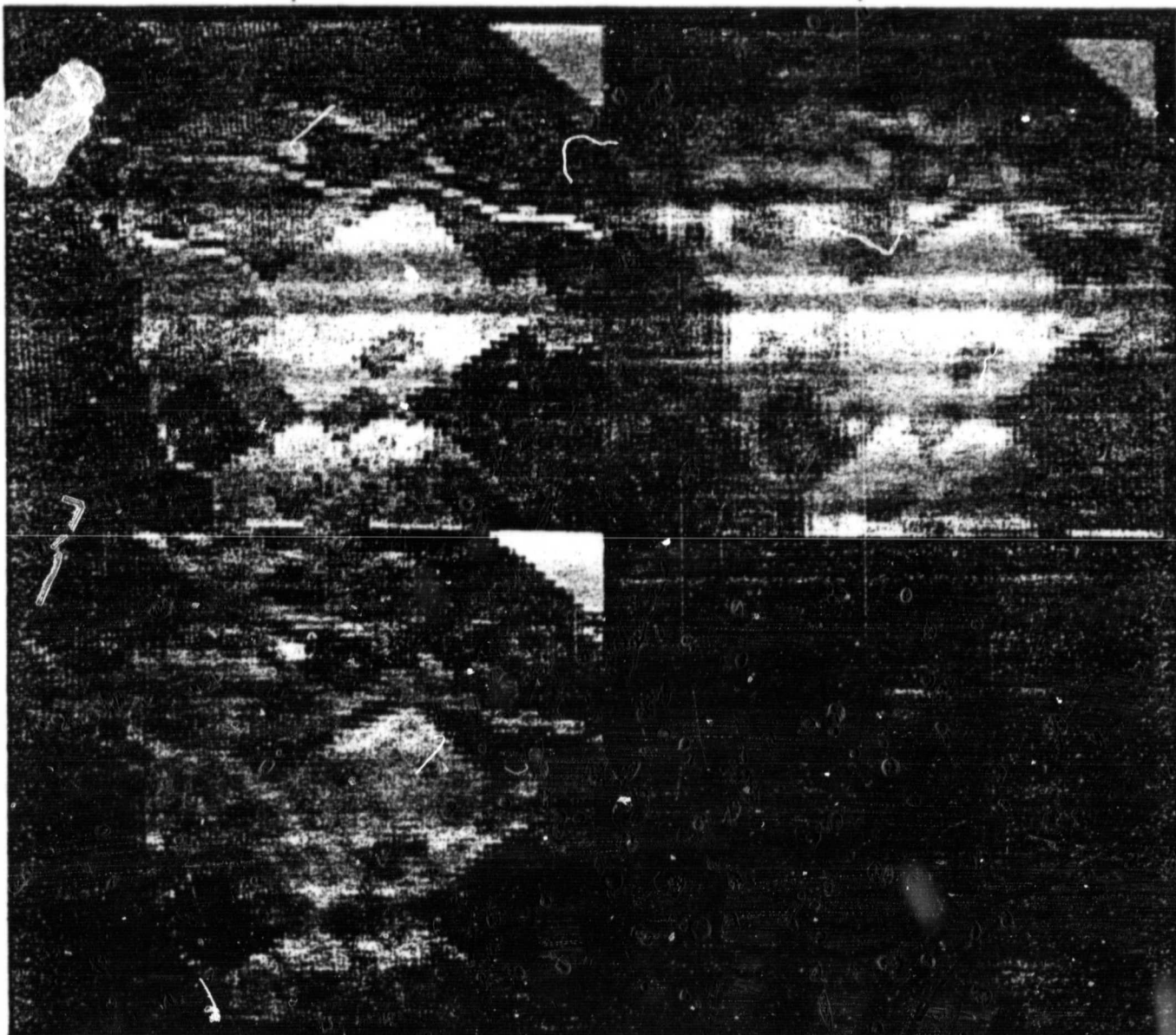
810344-113

a) ORIGINAL

$H_1 = 4.2$

b) INVERSE FFT

ERROR = 3.9
$H_1 = 4.3$

810344-114

ERROR = 3.7
$H_1 = 4.3$

c) MAXIMUM ENTROPY

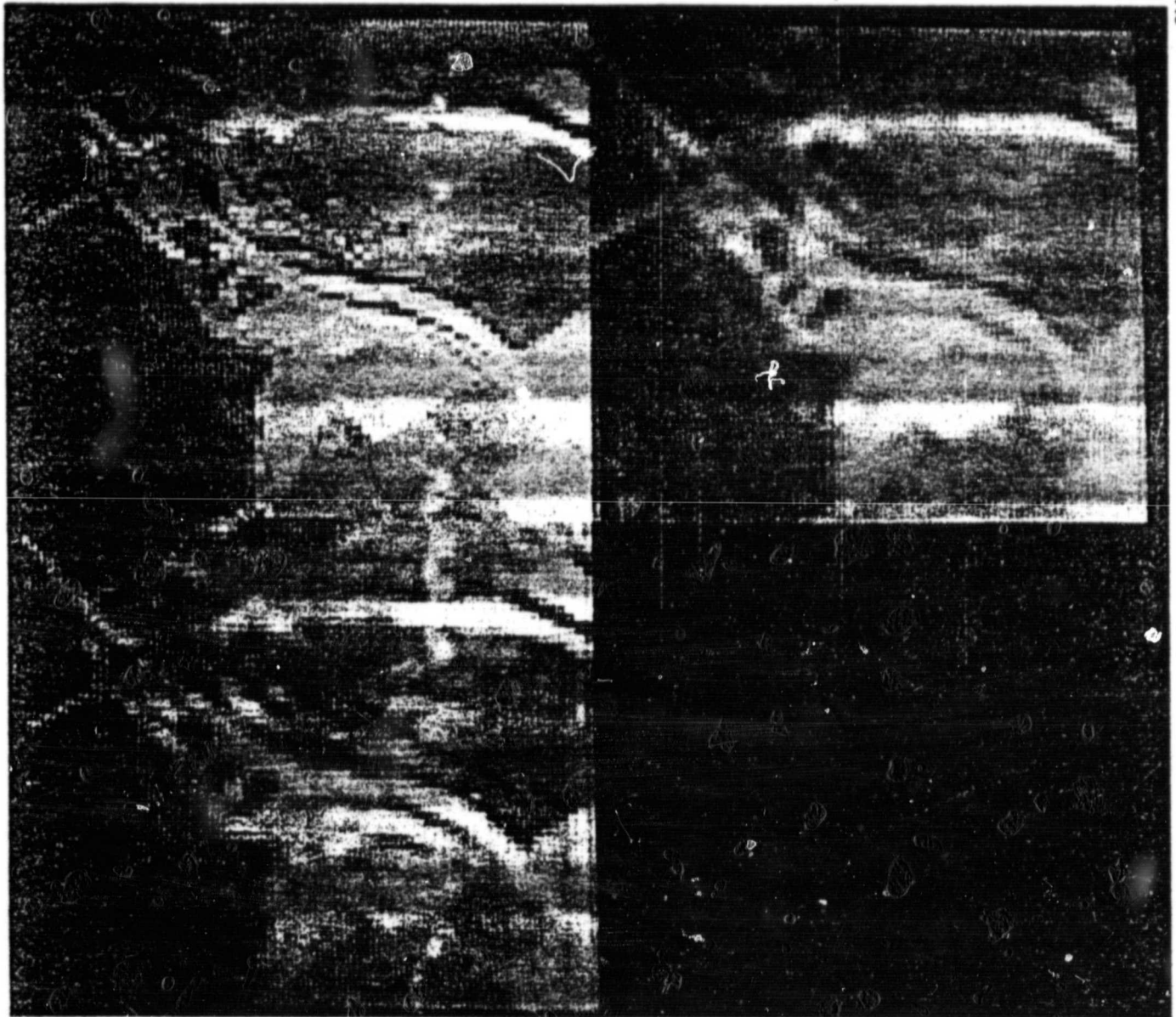FIGURE 4-8. SAMPLING FUNCTION S2 APPLIED ON IMAGE 2

C-3

a) ORIGINAL

$H_1 = 4.2$

b) INVERSE FFT

ERROR = 4.2
$H_1 = 4.2$

810344-115

ERROR = 4.0
$H_1 = 4.2$

c) MAXIMUM ENTROPY

FIGURE 4-9. SAMPLING FUNCTION S2 APPLIED ON IMAGE 3

a) ORIGINAL

$H_1 = 4.4$

b) INVERSE FFT

ERROR = 8.1
$H_1 = 4.3$

810344-116

ERROR = 9.2
$H_1 = 4.2$

c) MAXIMUM ENTROPY

FIGURE 4-10. SAMPLING FUNCTION S3 APPLIED ON IMAGE 1
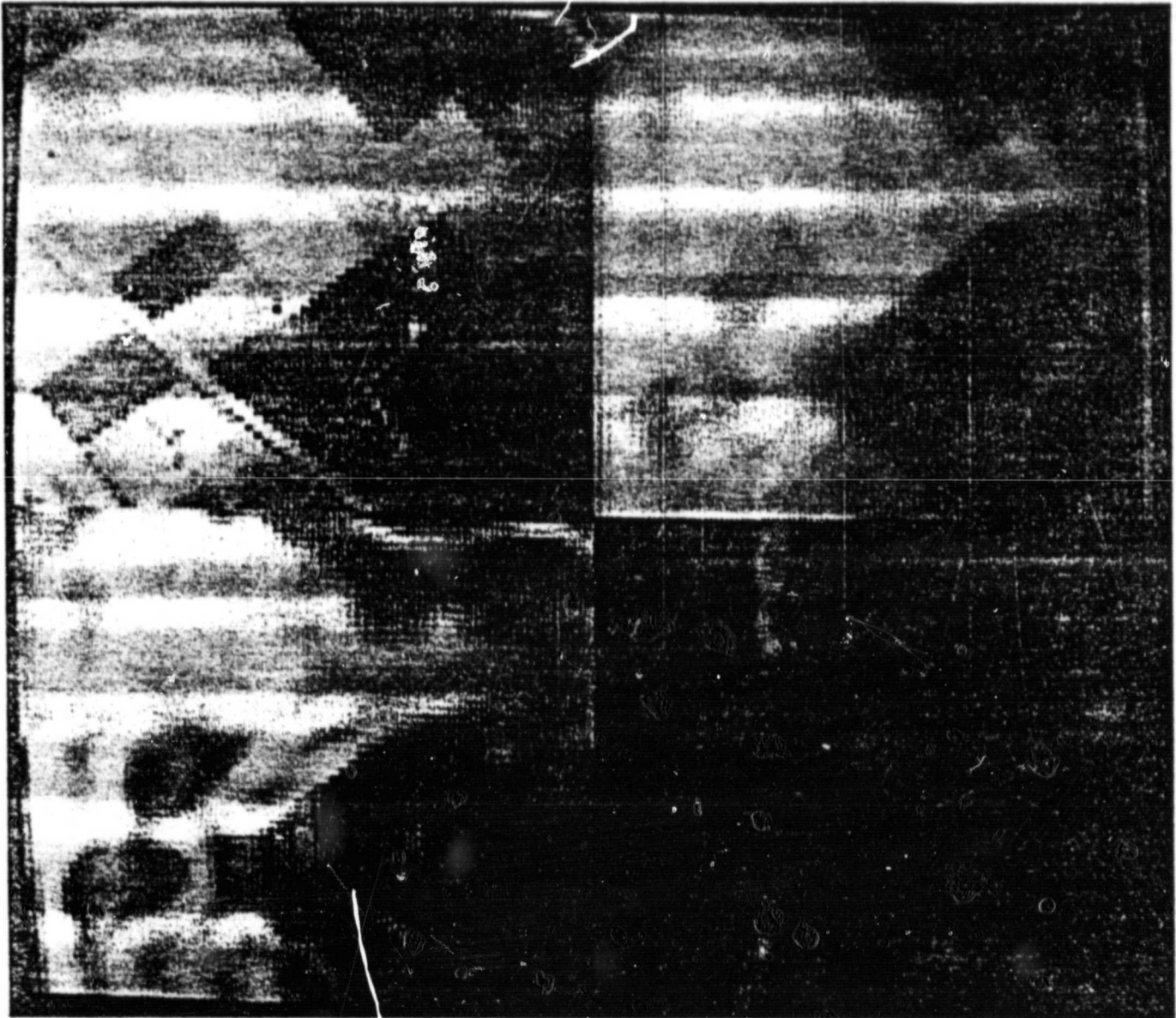
a) ORIGINAL

$H_1 = 4.2$

b) INVERSE FFT

ERROR = 6.5
$H_1 = 4.3$

810344-117

ERROR = 9.1
$H_1 = 4.0$

c) MAXIMUM ENTROPY

FIGURE 4-11. SAMPLING FUNCTION S3 APPLIED ON IMAGE 2
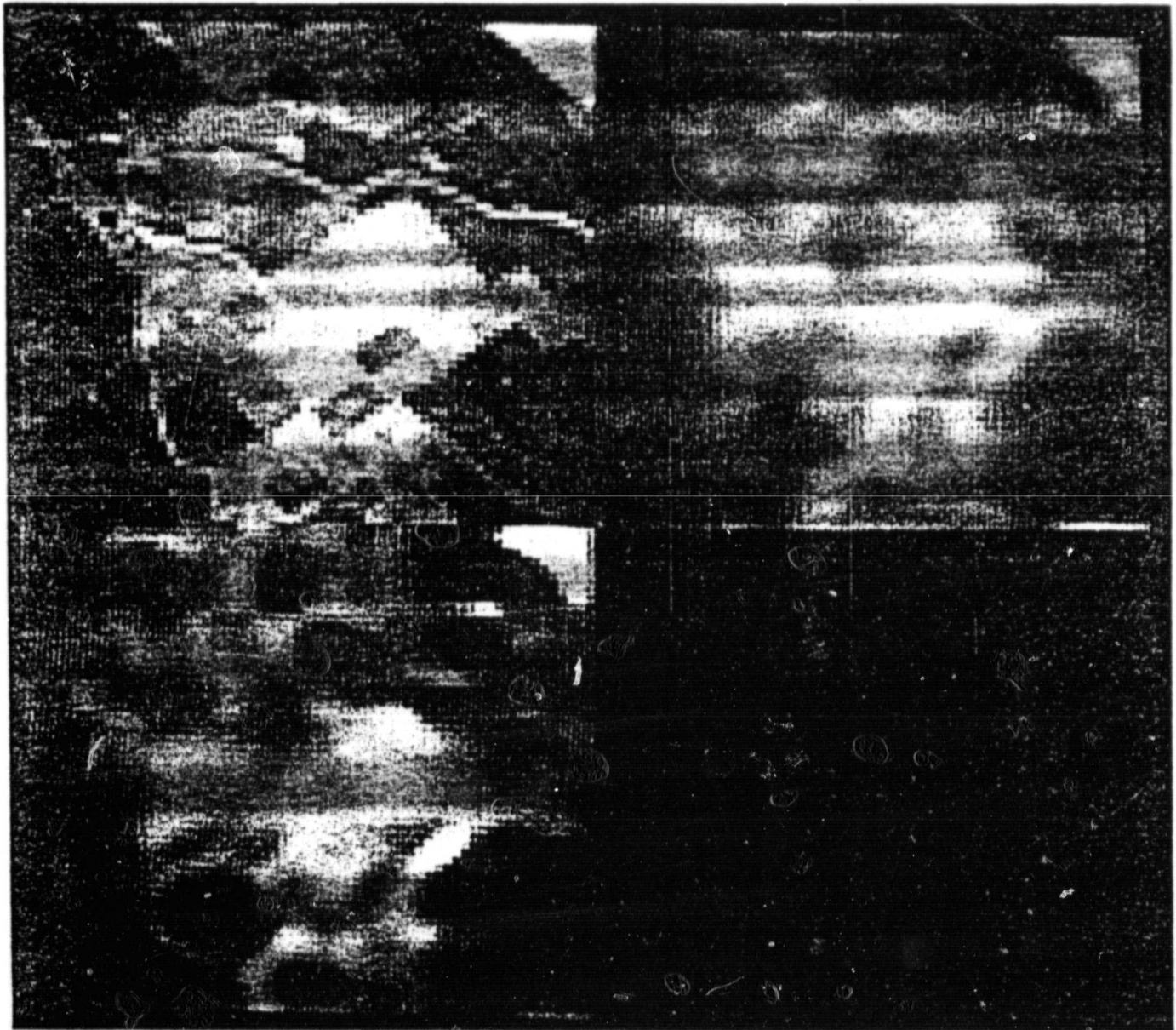
a) ORIGINAL

$H_1 = 4.2$

b) INVERSE FFT

ERROR = 6.9
$H_1 = 4.2$

810344-118

ERROR = 8.4
$H_1 = 4.1$
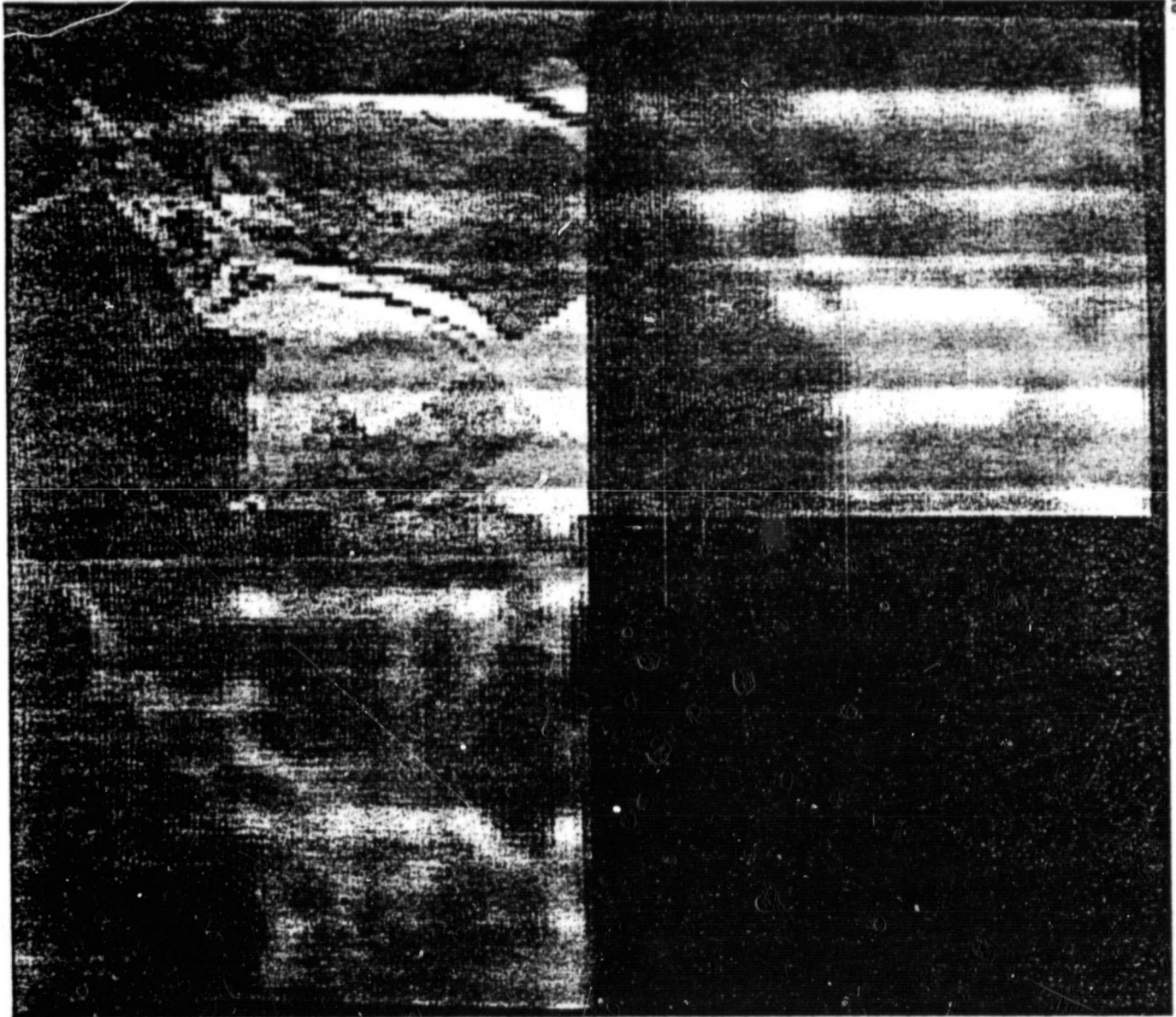
c) MAXIMUM ENTROPY

FIGURE 4-12. SAMPLING FUNCTION S3 APPLIED ON IMAGE 3

a) ORIGINAL

IMAGE 1

b) S2

c) S1

d) S3

FIGURE 4-13. DIRECT FFT INVERSE WITH VARIOUS SAMPLING FUNCTIONS

a) ORIGINAL

IMAGE 1

b) S2

810344-120
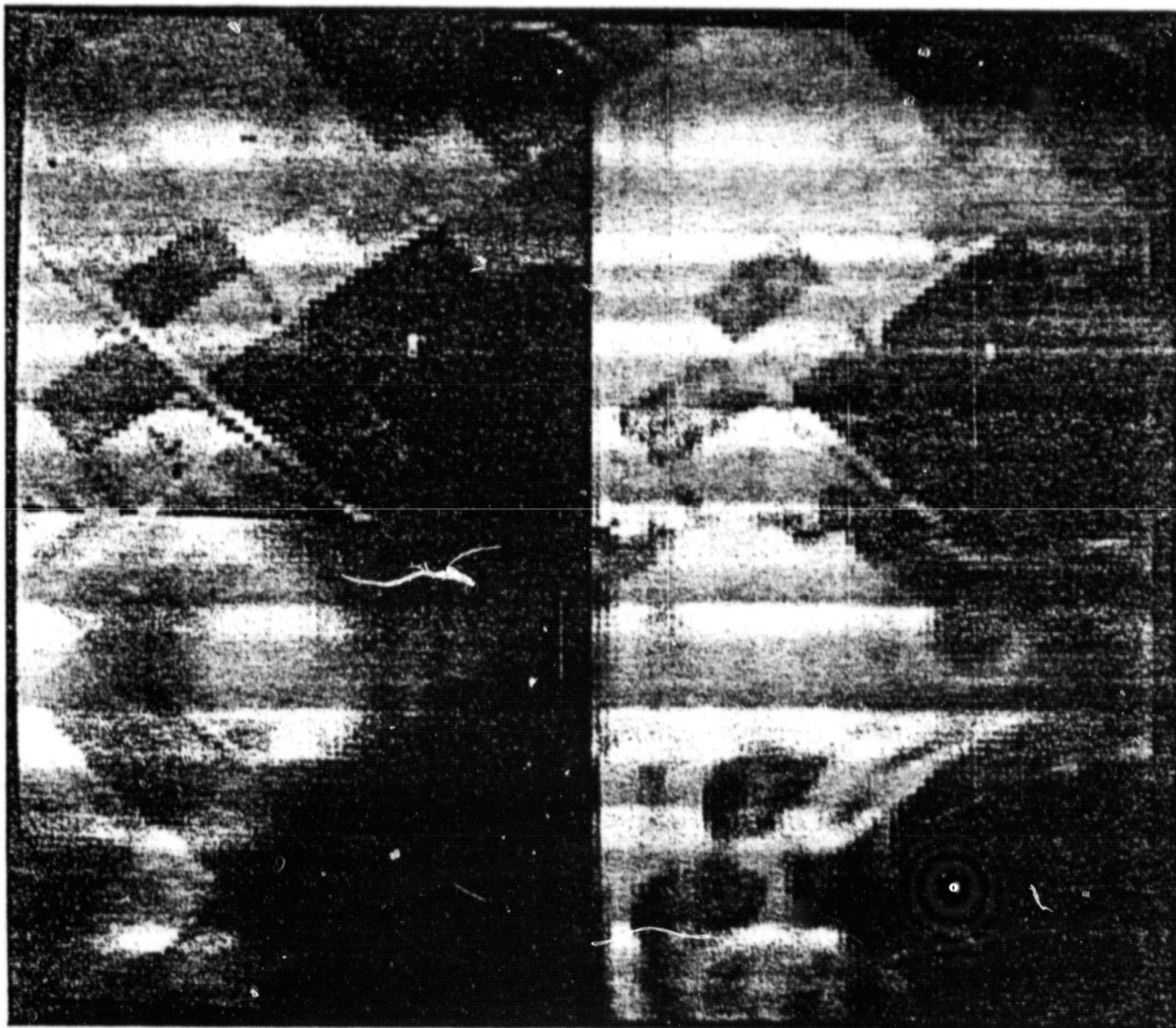
c) S1

d) S3

FIGURE 4-14. MAXIMUM ENTROPY WITH VARIOUS SAMPLING FUNCTIONS

TABLE 4-1. SQUARED-ERROR SUMMARY WITH TEST IMAGES

		S1	S2	S3
IMAGE 1	MEM	16.7	4.2	9.2
	FFT^{-1}	14.5	4.8	8.1
IMAGE 2	MEM	15.2	3.7	9.1
	FFT^{-1}	9.8	3.9	6.5
IMAGE 3	MEM	10.9	4.0	8.4
	FFT^{-1}	9.9	4.2	6.9

TABLE 4-2. CONVERGENCE TEST FOR MEM RECONSTRUCTION,
TEST IMAGE 3 AND TEST CASE 1, SAMPLING FUNCTION S1

	Test Image 3		Test Case 1	
	Squared-Error	Entropy H_1	Squared-Error	Entropy H_1
Steps where meaningful images were produced	12.13	3.09	36.9	1.8
	11.30	3.55	22.5	2.3
	11.10	3.65	15.1	2.2
	11.04	3.72	12.4	2.2
	10.91	3.83	10.7	2.2
	10.89	3.89	9.9	2.1
	10.88	3.92	9.7	2.1
	10.87	4.00	8.5	2.1
			8.3	1.9
			8.0	2.0
			7.4	1.9

TABLE 4-3. CONVERGENCE TEST FOR MEM RECONSTRUCTION,
TEST IMAGE 3, FULL SAMPLE

	Squared-Error	Entropy H_1
	4.14	4.27
Steps where meaningful images were produced	0.09	4.22
	0.02	4.23
	0.02	4.23
	0.02	4.23

5. CONCLUSION

This report is by no means a complete summary of image restoration
techniques. Several topics in the initial outline of work have not been
incorporated in the report due to lack of manpower and time. The high reso-
lution radiometer images from China Lake arrived too late to be of use in
this computer study. Analysis of antenna MTF and sensor PSF was not carried
out, although it is noted in Chapter 2 and Reference 1 that such linear
effects, if nonsingular, can be removed effectively by inverse filtering
technique for noise-free data, and Wiener filtering for noisy data.

In general, undersampling is quite undesirable. Many algorithms have
been developed to compress image data; some of them achieve very high com-
pression ratio for the particular class of images that they are designed for.
In all cases, the statistical properties of the images are studied carefully
and algorithms designed to remove redundancy as the rate distortion theory
suggests. One cannot expect to recover the original source resolution when
useful information is arbitrarily discarded. This is clearly illustrated by
the two sampling strategies where one is simulated multi-λ system and the
other is lowpass filtering system. Both systems have approximately the same
number of data samples. Given that all three test images have their spatial
frequency contents concentrated around zero frequency, the latter preserve
much more useful information than the multi-λ system and enables a better
reconstruction.

Maximum entropy reconstruction method is studied to combat the under-
sampling problem. It worked very well on the stacked cube test patterns,
even with simulated multi-λ sampling of frequency axes and diagonals. It
only offers a slight or no improvement over direct transform technique on
the three test radiometer images. Since the algorithm has builtin measure-
ment error and noise constraint, it may be a better reconstruction technique,
even the system is not undersampled. Further study is needed to verify the
universal worthiness of the method.

6. BIBLIOGRAPHY FOR PART II

(1) H.C. Andrew and B.R. Hunt, *Digital Image Restoration*, Prentice-Hall, New Jersey, 1977.

(2) T. Berger, *Rate Distortion Theory*, Prentice-Hall, New Jersey, 1971.

(3) A.J. Viterbi and J.K. Omura, *Principles of Digital Communication and Coding*, McGraw-Hill, New York, 1979.

(4) S.E. Roucos and D.G. Childers, "A Two-Dimensional Maximum Entropy Spectral Estimator," *IEEE Transactions on Information Theory*, vol. IT-26, pp. 554-560, Sept. 1980.

(5) A.K. Jain, "Image Data Compression: A Review," *Proc. of IEEE*, vol. 69, pp. 349-389, March 1981.

(6) A.N. Netravali and J.O. Limb, "Picture Coding: A Review," *Proc. of IEEE*, vol. 68, pp. 366-406, March 1980.

(7) A. Habibi (ed.), "Special Issue on Image Bandwidth Compression," *IEEE Transactions on Communication*, vol. COM-25, pp. 1249-1440, Nov. 1977.

(8) S.J. Wernecke, "Maximum Entropy Techniques for Image Reconstruction from Interferometer Measurements and Projections," Ph.D. Dissertation, Standford University, Stanford, 1976.

(9) S.J. Wernecke, "Two-Dimensional Maximum Entropy Reconstruction of Radio Brightness," *Radio Science*, vol. 12, pp. 831-844, Sept. 1977.

(10) S.J. Wernecke and L.R. D'Addario, "Maximum Entropy Image Reconstruction," *IEEE Transactions on Computers*, vol. C-26, pp. 351-364, April 1977.

(11) S.J. Wernecke and C.J. Grebenkemper, "An Iterative Program For Maximum Entropy Fourier Synthesis," Stanford Radio Astronomy Institute Technical Memo #582, Stanford, 1976.

(12) A.R. Thompson and R.N. Bracewell, "Interpolation and Fourier Transformation of Fringe Visibilities," The Astronomical Journal, vol. 79, pp. 11-24, Jan. 1974.

(13) H. Stark, "Sampling Theorem in Polar Coordinates," J. Opt. Soc. Am., vol. 69, pp. 1519-1525, Nov. 1979.

(14) F.A. Hagen, private communication.

(15) C. DeBoor, "Bicubic Spline Interpolation," J. Math. and Physics, vol. XLI, no. 3, Sept. 1962.

(16) M.S. Bartlett, An Introduction to Stochastic Processes, 2nd ed., Cambridge University Press, London, 1966.

(17) D.G. Childers (ed.), Modern Spectrum Analysis, IEEE Press, New York, 1978.

(18) W.K. Pratt, Digital Image Processing, John-Wiley & Sons, New York, 1978.