This dissertation has been 65-5631
microfilmed exactly as received

EBERLE, Jon William, 1934-

AN ADAPTIVELY PHASED, FOUR-ELEMENT
ARRAY OF THIRTY-FOOT PARABOLIC RE-
FLECTORS FOR PASSIVE (ECHO) COMMUNI-
CATION SYSTEMS.

The Ohio State University, Ph.D., 1964
Engineering, electrical

University Microfilms, Inc., Ann Arbor, Michigan

AN ADAPTIVELY PHASED, FOUR-ELEMENT ARRAY OF
THIRTY-FOOT PARABOLIC REFLECTORS FOR
PASSIVE (ECHO) COMMUNICATION SYSTEMS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Jon William Eberle, B. S. E. E., M. S. E. E.

The Ohio State University
1964

Approved by

Adviser

Department of Electrical
Engineering

ACKNOWLEDGMENT

The author would like to express sincere appreciation to his adviser Professor C. H. Walter and to Professors W. H. Peake and J. H. Richmond for their critical evaluation of the manuscript and for freely giving of their time for consultation and suggestions.

Secondly the author would like to express sincere appreciation to his colleague Mr. Stephen L. Zolnay for unselfishly giving countless hours of his time throughout the three years that was required for the construction, checkout and operation of the array. During these three years, there were untold problems and discouragements and if it had not been for his persistence and determination, the present operational status of the array could not have been achieved. Additionally his assistance in the analysis of the data is gratefully acknowledged and it should be pointed out that Chapter V and Sections D and E of Chapter VI have been excerpted from technical reports coauthored with him.

Several others have been of great assistance in bringing the array to an operational status and have also freely given of their time. The author is especially grateful to Delmer Hayes,

ACKNOWLEDGMENT -- Continued

James Tourdot, Brenton Groves, Ronald Huff, Dean Svoboda and to Frank Cook, Bob Evans, Max Gordon and Dave Landis. It it had not been for the excellent work and perseverance of these individuals, the construction and instrumentation of the array would have been impossible. Additionally the unfailing assistance of Frank Cook, Don Henry and Karl Shaffer in the analysis of data is gratefully acknowledged.

The continuing support under Contract AF 30(602)-2166 from Rome Air Development Center and particularly the splendid cooperation of Mr. James E. McFee and Mr. James W. Bailey is sincerely appreciated.

The author gratefully acknowledges the expert work of Mr. Lew Kail and Max Katzenbach in drafting, Mrs. Barbara Kerwood in typing and Messrs. Bill Ryan and Jim Puckett in printing. Also the expert editing by Mr. Chester E. Ball is gratefully acknowledged.

The author is indebted to Professor C. A. Levis and Mr. R. A. Fouty for minimizing his duties in the Antenna Laboratory during the time in which the manuscript was prepared. If it had not been for this assistance on their part, the preparation of the

ACKNOWLEDGMENT -- Continued

manuscript would have been much more difficult and would have taken much longer. Also their splendid cooperation on administrative matters throughout the course of the work is gratefully acknowledged.

Last, but by no means least, the author is indebted to his wife and children, who have made many sacrifices during this educational endeavor.

VITA

August 28, 1934

Born — Chillicothe, Ohio

1954..... Technical Assistant, Antenna Laboratory,
Department of Electrical Engineering,
The Ohio State University, Columbus, Ohio

1957..... B.S. E. E., The Ohio State University,
Columbus, Ohio

Research Associate, Antenna Laboratory,
Department of Electrical Engineering,
The Ohio State University, Columbus, Ohio

1960..... M.S. E. E., The Ohio State University,
Columbus, Ohio

1962..... Assistant Supervisor, Antenna Laboratory,
Department of Electrical Engineering,
The Ohio State University, Columbus, Ohio

CONTENTS

	Page
ACKNOWLEDGMENTS.....	ii
VITA.....	v
TABLES.....	ix
ILLUSTRATIONS.....	x
Chapter	
I INTRODUCTION.....	1
II INITIAL CONSIDERATIONS IN THE DESIGN OF A PHASED ARRAY.....	9
A. Phase Perturbances	9
1. Base line changes	9
2. Differential doppler	11
3. Phase scintillations	14
4. Feedline effects	14
5. Equipment phase jitter	15
B. Signal Distribution in the Array	16
C. Array Bandwidth	19
D. Array Pattern Control	21
III DESCRIPTION OF ARRAY AND DATA PROCESSING SYSTEM:.....	22
A. Introduction	22
B. Overall System	23
1. General	23
2. Monopulse tracking system	27
3. Signal acquisition system	30
4. Signal tracking and cohering equipment	35

CONTENTS -- Continued

Chapter		Page
IV	DETAILED SYSTEM ANALYSIS.....	44
	A. Introduction	44
	B. Linearization	45
	C. Root Locus Method of Analysis	48
	D. Bode Plot	58
	E. Transient Response	61
	F. Error Signal Considerations	65
	G. Effect of Signal Strength on Bandwidth	68
	H. Tracking Range	73
	I. Noise Effects	75
V	SIGNAL TO NOISE RATIO IMPROVEMENT.....	80
	A. Introduction	80
	B. Discussion	82
	1. Summing receivers	82
	2. Measurement using RMS voltmeter	85
	3. Measuring using probability density function	89
	4. Measurement to demonstrate signal coherence and noise incoherence	92
VI	DATA ANALYSIS AND INTERPRETATION OF ECHO II REFLECTED SIGNALS.....	96
	A. Echo II Description	96
	B. Stationary Characteristics of Echo II Data	99
	1. Introduction	99
	2. Discussion	100
	3. Data	106
	4. Conclusions	123
	C. Probability Density Functions of Echo II Reflected Signals	128
	1. Introduction	128
	2. Results	134
	3. Conclusions	148

CONTENTS -- Continued

Chapter	Page
D. Short-Term Autocorrelation Function	149
1. Introduction	149
2. Discussion	150
a. Specular-to-scattered power ratio	151
b. Inverse transform	154
c. Surface roughness	158
d. Discrete frequencies	161
e. Fading frequencies	162
E. Long-Term Autocorrelation Function	165
1. Introduction	165
2. Discussion	165
3. Data	168
4. Conclusions	170
VII CONCLUSIONS.....	171
APPENDIX	
A COMPOUND INTERFEROMETER.....	174
B WIDEBAND INTERFEROMETER.....	181
C OPTIMUM GAIN.....	188
D MAXIMUM SNR WITH OPTIMUM GAIN.....	190
E PHASE-LOCKED DEMODULATOR.....	191
F MEASUREMENT OF SNR.....	194
1. Introduction	194
2. Conventional Measuring Instruments	195
3. Random Waveform Statistics	197
4. Measurement Technique Using an Oscilloscope	199
BIBLIOGRAPHY.....	206

TABLES

Table		Page
I	The decrement factor ζ versus the bandwidth and the peak overshoot ratio (POR)	58
II	Specular to scattered power ratio for selected Echo II revolutions	143
III	Range of fading for selected Echo II revolutions	145

ILLUSTRATIONS

Figure		Page
1	Elementary two element array.....	10
2	Feedline attenuation effects.....	17
3	Phase relationship for a mixer.....	18
4	Four element phased array.....	24
5	Feed system block diagram.....	26
6	Monopulse tracking receiver.....	28
7	Complex vector sum used for monopulse tracking error signal.....	29
8	Acquisition equipment.....	32
9	Fan beam geometry.....	33
10	Signal cohering and combining receiver.....	34
11	Local oscillator.....	36
12	Element and array antenna pattern.....	39
13	Phase locked demodulator with quadrature detector.....	40
14	Overall view of mount control, cohering receiver and local oscillator.....	42
15	Overall view of data recording system.....	42
16	Overall view of data analysis system.....	43

ILLUSTRATIONS -- Continued

Figure		Page
17	Linearization of the phase locked loop.....	48
18	Measurement of angles for the root-locus.....	50
19	Typical root-locus for phase locked demodulator.....	53
20	Definition of damping factor, undamped natural frequency damped natural frequency and decrement factor.....	55
21	Step function response measurement.....	57
22	Root-locus with open loop gain.....	70
23	Bode plot of phase locked demodulator with frequency input.....	72
24	Bode plot of phase locked demodulator with phase input.....	72
25	Probability density function of the phase of a sinusoid plus Gaussian noise.....	78
26	Combination of unequal SNR's.....	83
27	Summing of unequal SNR in a four channel array.....	87
28	SNR improvement, Mode I, frequency controlled local oscillator.....	90
29	SNR improvement, Mode I, phase controlled local oscillator.....	90
30	SNR improvement, Mode II, frequency controlled local oscillator.....	91
31	SNR improvement, Mode II, phase controlled local oscillator.....	91

ILLUSTRATIONS -- Continued

Figure		Page
32	SNR improvement, Mode III, phase controlled local oscillator.....	92
33	SNR improvement from probability density function.....	93
34	Signal coherence from 28 Mc phase detector recordings.....	95
35	Ground inflation of Echo II.....	97
36	Orbital inflation of Echo II.....	98
37	Time and ensemble averages.....	101
38	Convergence of the first probability density function of pulses for Echo II revolution 407....	109
39	Convergence of the first probability density functions of envelope between pulses for Echo II revolution 407.....	110
40	Convergence of the first probability density functions of Echo II revolution 1730, CW modulation.....	111
41	Convergence of the first probability density functions of Echo II revolution 1730, CW modulation, 30 samples per second.....	112
42	Convergence of the first probability density functions of Echo II revolution 1730, CW modulation, 50 samples per second.....	113
43	First probability density function for consecutive 30 second time periods for Echo II revolution 1730.....	115

ILLUSTRATIONS -- Continued

Figure		Page
44	First probability density function for consecutive 30 second time periods for Echo II revolution 1730, 50 samples per second.....	116
45	First probability density function for consecutive 30 second time periods for Echo II revolution 1901, 50 samples per second.....	118
46	First probability density function from ensemble of records of Echo II revolution 1730.....	119
47	Convergence of autocorrelation function for Echo II revolution 407	121
48	Convergence of autocorrelation function for Echo II revolution 1730.....	122
49	Autocorrelation function for Echo II revolution 1730 for consecutive 30 second time periods.....	124
50	Autocorrelation function for Echo II revolution 1901 for consecutive 30 second time periods.....	125
51	Oscillographic display of received pulses.....	129
52	Coordinate system for probability.....	135
53	Probability density functions of common waveforms.....	137
54	Probability density functions of selected Echo II revolutions.....	140
55	Specular to scattered power ratio.....	143

ILLUSTRATIONS -- Continued

Figure		Page
56	Probability density function for Echo I.....	146
57	Definition of specular to scattered power, ratio from the autocorrelation function.....	153
58	Effect of the coincidence of the axis of rotation and the angle of look.....	157
59	Long term autocorrelation function for Echo II revolution 1210.....	169
60	Long term autocorrelation function for Echo II revolution 1901.....	169
61	Simple interferometer with phase shifter.....	175
62	Effects of adding interferometer elements in a compound interferometer.....	178
63	Examples of far-field patterns and aperture autocorrelation functions for three aperture distributions.....	184
64	Aperture autocorrelation function and power pattern of a two-element interferometer receiving signals over a wide band of frequencies.....	185
65	Aperture autocorrelation function and aperture distribution for a four element interferometer.....	186
66	Rayleigh density function.....	199
67	Probability density function of envelope of sinusoid plus noise.....	200
68	Oscilloscope display of sinusoid plus noise.....	202
69	Signal to noise ratio versus signal plus noise to noise ratio.....	205

CHAPTER I INTRODUCTION

With the advent of space communications, the need for large antenna apertures has increased tremendously. This need is not in the same sense as the need for large apertures in radio astronomy. In many cases the radio astronomer has sufficient signal with which to work, either inherently or by use of special detection techniques, but he frequently does not have sufficient resolution for position determination. Although this situation also exists in satellite work where position determination is needed to establish the orbit of a satellite or its orbital elements, the most often encountered situation is one in which resolution is of lesser importance than antenna gain. This type of situation exists when broad information bandwidths are involved at very large distances. For a given signal-to-noise ratio needed for operation above the threshold of a demodulator and for a given amount of signal power and a given system noise temperature, the bandwidth of a communication system is inversely proportional to the square of the range for one-way transmission. There are several techniques which may be used to increase the bandwidth for a given range. These are listed as:

1. The required signal-to-noise ratio for demodulation can be decreased by use of more sophisticated demodulators, such as phase-locked demodulators, frequency-locked demodulators,

maximum likelihood demodulators[1], etc. Also more sophisticated types of modulation can be used to aid in the demodulation process.

2. The signal power available for the channel can be increased. This approach is limited for on-board transmitters because of weight limitations, and for ground transmitters because of economies and state of the art.

3. The system noise in the receiving system can be reduced by proper antenna design and by use of low-noise amplifiers, such as parametric amplifiers and masers.

When particular attention has been directed toward each of these approaches, it is found in many instances that the available signal-to-noise ratio in a desired bandwidth is not sufficient for a given range. Two alternatives then exist: (a) the bandwidth may be reduced to yield sufficient signal-to-noise ratio, or (b) the signal may be increased by the use of higher antenna gain. This paper is directed toward the latter approach, i. e., that of higher antenna gain.

The gain of an antenna system is directly proportional to the aperture area of the antenna, whereas resolution is dependent upon the way the antenna signals are processed. For example, resolution may be increased by using interferometer techniques where resolution is proportional to the spacing between elements of the interferometer, but the gain of the interferometer is proportional to the aperture area

of the interferometer elements and neglecting coupling, independent of the spacing between them. Thus, antenna gain can be increased only by increasing the aperture area of the antenna system. This approach has been evidenced in recent years by the larger and larger sizes of single aperture antennas. Up to the present time steerable apertures as large as 300 feet in diameter have been constructed; and one antenna, which is fixed in position, has been constructed which is 1000 feet in diameter. This approach presents some very real problems when the aperture sizes become larger than 100 feet in diameter and when the further requirements of high-frequency operation and wide beam coverage are imposed.

One of the most serious problems in this approach lies in constructing such large apertures to the close mechanical tolerances which are required for higher frequency operation. Antenna surface accuracies on the order of one-sixteenth of the operating wavelength must be maintained to achieve nominal gain from the antenna. As the antenna size is increased and the operating wavelength is decreased, a point is reached where the state-of-the-art will not permit further increases in size or decreases in wavelength. This situation may be likened to the gain-bandwidth product limitations of an amplifier, in which the analogous product would be the antenna diameter-operating frequency product. At the present time an antenna 110 feet in diameter

and operating at frequencies up to 15 GC is in operation; this is considered state-of-the-art for antenna diameter-operating frequency product. It appears that attempts to produce an antenna larger than this for higher frequencies would involve problems with the modulus of elasticity of commonly used materials, such as steel or aluminum.

In addition to the mechanical problems involved with the construction of the antenna, similar problems are experienced with the design and construction of the mount which would be used to steer large antennas. As the size of the antenna is increased, the beamwidth of the antenna decreases proportionally, thereby requiring that the large antenna be steered or pointed to higher accuracies. This is a situation which quickly becomes extremely expensive, cumbersome, and ultimately impractical. Also associated with the decrease in beamwidth which results from increasing antenna size are problems of an electrical nature. One is the problem of acquiring a source of signal or a target with a narrow antenna beamwidth, particularly when the target or signal source location is not precisely known. This situation calls for some type of beam-scanning, which is customarily accomplished by mechanically sweeping the antenna in an interlace or spiral scan over the solid angle in which the signal source or target is expected. Thus scanning imposes large mechanical stresses on the antenna mount. Another problem is position scintillation of a source

of signal due to atmospheric and/or ionospheric effects when operation is near the horizon[2]. Thus scintillations, which can become noticeable and troublesome when beamwidths on the order of 1 minute of arc are involved, are of a random nature and would require some type of beam steering, which would, in all likelihood, involve antenna motion.

The preceding paragraphs outlined several of the disadvantages and problems which apply to the approach of using single large apertures to achieve high antenna gain. The use of phased-arrays is a partial answer to each of the disadvantages or problems. The concept of the phased-array involves the use of small apertures to achieve the equivalent performance of a single larger aperture. Because smaller antennas are used, they can be made more accurately and at lower cost. Because of the smaller size, the beamwidths of the individual elements are larger and hence do not require as large or as accurate an antenna mount as does the single large aperture. Thus a considerable saving in original equipment costs is effected. With regard to the electrical type problems associated with single, large apertures, the phased-array approach offers considerable flexibility in the acquisition process; i.e., several individual antenna beams are involved and they can be caused to overlap in such a way that a larger beamwidth is available for initial acquisition. Also, since the beams of the phased array

can be electrically controlled, random scintillations in the position of a signal source or target can be followed by electronic rather than mechanical means.

It should be pointed out that there are several types of antennas which are classified as phased arrays. Many of the phased arrays in existence today consist of numerous small elements which are combined by means of directional couplers, making a dispersive type system. Thus by changing the operating frequency, the beams can be caused to scan[3] . Another type of phased array is similar to the type mentioned above except for the way in which the feed lines are arranged. By making all feedlines the same electrical length, a non-dispersive system is obtained in which the positions of the beams are somewhat independent of frequency[3] . In a third type of phased array, exemplified by the ESAR[4] (Electronically Steerable Array Radar) and the SPADATS[5] (Space Detection and Tracking System) system, numerous elements are excited by a computer-controlled tapped delay line. Through the use of computers in these systems, the simultaneous functions of tracking several targets and scanning for new acquisitions is possible. These concepts have been proved by operational systems. In a fourth type phased array, with which this paper is concerned, the individual elements of the array are, in themselves, large antennas which are independently controlled or steered. The elements in the particular array of concern are parabolic reflectors 30 feet in diameter;

but for other situations, larger or smaller array elements might be used. The spacing between elements depends upon the particular application of the array, but the spacing can vary from twice to several times the element diameters.

Although the phased array consisting of large, independently steered antenna elements offers a number of advantages for achieving high antenna gain, it should not be considered as a panacea for all high-gain antenna problems. There are several areas in which the phased array present some difficult problems when compared with the single large aperture. Many of these problems are associated with the phased array's multiple-beam characteristics which are caused by the wide element spacing. When a single beam is desired, such as for transmission applications or for operation in a multiple target situation, the array approach is at a distinct disadvantage compared with the single-aperture approach. Also, for situations in which the gain of the entire aperture is required for tracking functions, the array approach presents problems. These problems are currently under study and, in time, solutions may be found which would make the array more usable in such circumstances. At the present time, the array approach considered here has been specifically studied for receiving applications; it can be said for this type of operation that the array approach offers a trade off between the mechanical complexity of the single aperture and

and the electrical complexity of the phased array. This trade off must be carefully studied before a decision can be made as to what type of antenna would be used in a given situation.

CHAPTER II

INITIAL CONSIDERATIONS IN DESIGN OF A PHASED ARRAY

A. Phase perturbances

There are several important problem areas which must be carefully considered in the design of an array of independently steered large antenna elements. Perhaps the most important of these problems is the types of effects for which corrections must be made in attempting to maintain the signals from all of the antenna elements in phase at a common summing point. In fact, this is one of the underlying principles in the operation of a phased array; that is, to obtain an increase in signal-to-noise ratio at the output of a summing receiver, the signals should be added coherently (in-phase) while the system noise adds incoherently. Several effects which cause the phases of the signals to vary between antenna elements are detailed below.

1. Base line changes

Consider a two-element array as shown in Fig. 1. The spacing between the elements is d and the element patterns are both given by $E(\theta)$. From the figure, it can be seen that the phase between the two elements is [6]

$$(1) \quad \psi = Kd \sin \theta,$$

Fig. 1--Elementary two element array.

and that as θ (the angle of arrival of the received signal) varies, the phase will also vary. The antenna voltage pattern of the two-element array is given by

$$(2) \quad P(\theta) = 2E(\theta) \cos \left[\left(\frac{\pi d}{\lambda} \right) \sin \theta \right],$$

which is a typical interferometer pattern. As the distance d increases, the phase changes more rapidly, thus causing the antennas to phase in and out more rapidly for a given change in θ , which results in more lobes in the antenna pattern. In a phased-array antenna, it is desirable to eliminate the lobed nature of the antenna pattern. To accomplish this, a phase-shifter would have to be inserted in one or both arms of the antenna to compensate for the phase shift given in Eq. (1). The rate of change of this phase shift is given by

$$(3) \quad \frac{d\psi}{d\theta} = Kd \cos \theta ,$$

where the maximum rate of change is required in a broadside direction and the minimum is in the endfire direction. It should also be noted that the phase-shifter, as well as its rate of change, must vary in a sinusoidal fashion.

When the phase-shifter is made to follow the variation given by Eq. (1), the effect on the antenna pattern is such that the interferometer lobes scan within the element pattern $E(\theta)$ in such a way that one of the interferometer lobes stays centered on the signal source. This simplified explanation serves to demonstrate the need for a phase-shifting system to compensate for the changes in baseline caused by scanning. The explanation also points out the need for a second type of tracking system which would keep the element pattern, $E(\theta)$, centered on the source of signal so that maximum signal can be obtained from each element. This second type of tracking involves mechanical movement of the array elements and can be accomplished by use of monopulse or conical scanning techniques.

2. Differential doppler

Several types of doppler frequency shifts must be considered in the design of a communication system involving high-velocity satellites or targets in conjunction with phased-array antennas[7]. The first type which must be considered in reception by a phased array or by

any receiving system will be referred to as overall or coherent doppler. This doppler is caused by overall radial motion of the signal source toward or away from a reference point and can be expressed as

$$(4) \quad f_A = f_o \frac{1}{1-(v/c)^2}$$

where f_A is the actual frequency after doppler shift. To good accuracy, the doppler frequency can be expressed by

$$(5) \quad f_d = f_o \frac{2v_r}{c}$$

where f_d is the doppler frequency shift, f_o is the transmitted frequency, v_r is the radial velocity, and c is the velocity of light. In the particular system of concern in this paper, the coherent doppler varied between ± 100 KC. Since all antennas in an array would experience this doppler, a single correction can be made to all the antennas of the array.

A second doppler effect in an antenna array for which corrections must be made is that referred to as differential doppler. This doppler is caused by a slightly different radial velocity component toward the individual elements in an antenna array. The very small parallax between elements of the array and the signal source is sufficient to cause the signal frequency received by each array element to be slightly different from the other array elements. When phases are to be compared, it is implied that all frequencies are equal, and hence to

compare and adjust the phase of elements in an array, the frequencies must all be the same at the point where phase is measured. In the system of concern here, the differential doppler frequency between array elements amounted to a maximum of 2 cps. This doppler component is directly dependent upon the operating frequency and upon the distance between array elements. Since phase is of concern between array elements to accomplish coherence and since frequency is proportional to the rate of change of phase, the problem of phase control between elements can be viewed as one of designing a control system which corrects for a velocity error; the velocity error being the cause of the differential doppler. This consideration is important in the choice of the type of control system, since the type of system will dictate the characteristics of the error signal. For example, a type O control system requires a constant error signal to maintain a constant value of the controlled variable; a type I system requires a constant error signal to maintain a constant rate of change of the controlled variable; and a type II system requires a constant error to maintain a constant acceleration of the controlled variable, where, in the problem at hand, the controlled variable is phase. Hence it is important in considering the differential doppler effects in a phased-array system. The magnitude of the differential doppler may appear small but its consideration is imperative and has a strong influence on the system design.

3. Phase scintillations

As noted in the introduction, there is a phase scintillation or a position scintillation on a signal source which lies beyond the atmosphere and/or the ionosphere. This scintillation is a function of the operating frequency and of the elevation angle of the source. It has been shown by Zolnay[8] that at a frequency of 2000 mc, the scintillations in position were less than 6 minutes of arc, as measured by an interferometer with a baseline of 60 feet. As the antenna size and the operating frequency increase, the antenna beamwidth could easily approach the 6-minute-of-arc value; in fact, it could be even smaller. If this situation were encountered, some means of correcting for the movement of the signal source would be required. Since the phase scintillations are of a random nature, they cannot be predicted, hence some technique is necessary to generate an error signal for this phase correction. For very large antennas, this is an important consideration in the design of the system.

4. Feedline effects

Whenever an rf signal is distributed over an appreciable distance by means of a transmission line, there are many possibilities for phase shift between the two points. A mere flexing of the cable is sufficient to cause a phase change of many degrees; and from this, compounded with heating and cooling effects, connector discontinuities,

moisture and ageing effects, a number of electrical degrees could be obtained through a transmission line. This is very important in a phased array in which the signals are brought from the individual elements to a common point for phase measurements and summing. Although the rate of change of phase through the cable may be rather slow, the amount of phase change may be large enough to require compensation when phasing the signals. Another effect, similar to the transmission line problem, is that of the shift of the phase centers of the antennas as the antennas are scanned. This is analogous to a change in the distance d in Fig. 1 as the angle θ is varied. In practice it is very difficult to make antennas, feed support spars, antenna mounts, etc. exactly alike; there is invariably some change in the distance between phase centers which should be accounted for in the design of an array.

5. Equipment phase jitter

In any control system, there is always some residual error present. This is true of closed loop control systems where a small error is needed to actuate the system. The amount of the error, or phase jitter, is dependent to a large extent upon the design of the control system and is an important consideration. Although this effect may be small in itself, it along with a combination of the above outlined effects may be cumulative so that a result well below theoretical capability may be obtained.

By considering each of the above perturbing influences, it can be concluded that a system with the ability to adapt itself to the various effects would be highly desirable. The other approach to the problem would be one of predicting the various perturbances and programming the corrections into the system by means of a computer. However, as pointed out, several of the effects are of a random nature which could not be accurately predicted. From these observations it was concluded that the desirable system would be a closed-loop control system which would adapt itself to changes in the phase between elements in an adaptive or self-correcting fashion.

B. Signal distribution in the array

Another problem which should be considered in the design of a phased array of the type being considered here is what would be the best way to route the signals from the antennas to the common phase-measuring and summing point. Two approaches which may be used are (1) the use of rf transmission lines[9] and (b) the use of IF lines. When rf lines are used, there are several detrimental factors which can greatly complicate the instrumentation of the array. One of these has been outlined above under feedline effects. If the RF is transmitted through cable or waveguide, the cable length need only change a small part of a wavelength to effect a phase shift of several degrees. More important than this, however, is the effect of the loss

in the cable on the array gain and its noise temperature. Since the purpose of the array is an increase in signal-to-noise ratio, one way to rate the array would be the ratio of its gain to its noise temperature[10]. Figure 2 shows one channel of a phased array from the antenna to the

Fig. 2--Feedline attenuation effects.

summing point where the signal is routed from the antenna by means of an RF transmission line having an attenuation α nepers/foot. If the antenna and amplifier noise temperature is T_a and the transmission line is at an ambient temperature T_o , then the noise temperature at the end of the feedline cable is given by

$$(6) \quad N_T = T_a e^{-\alpha l} + T_o(1 - e^{-\alpha l})$$

From this it may be seen that a compounding effect takes place, that is, the noise temperature and signal are each attenuated by $e^{-\alpha l}$, but an additional noise term, which amounts to $T_o(1 - e^{-\alpha l})$ is added to the system. By suitably choosing the gain of the amplifier in Fig. 2 so that the antenna noise is very much larger than noise from the feed line, the effect of the feedline could be minimized.

Another approach that can be used to distribute the signals in the array is to mix the incoming frequency down to an IF frequency and then distribute the IF signal over cable to the summing point[11] . Of concern in this process is the phase of the IF relative to the phase of the input signal. From Fig. 3 it can be seen that the phase of the IF signal of

Fig. 3--Phase relationship for a mixer.

frequency $(\omega_1 - \omega_2)$ is dependent upon the phase of the input signal, as well as that of the local oscillator. If the same local oscillator is used in each channel of the array, its phase would be present in all channels and hence could be considered as a reference at zero degrees. The advantage of this approach is that high gain is easily obtained at the IF frequency and that the cable losses at IF are much lower than those at RF. An additional advantage is that the wavelength at IF is much longer than at RF, hence a given change in cable length represents a much smaller phase shift at IF.

C. Array bandwidth

A third consideration in the design of the array is that of the inherent array bandwidth. Whenever the pattern of an interferometer is given, it is implied that the measurement was made at a monochromatic frequency. Since an interferometer is a dispersive aperture type antenna for other than broadside operation, the pattern of the interferometer is frequency dependent. It is on this basis that the instantaneous bandwidth of the array is defined[12]. Consider a simple two-element array as shown in Fig. 1, for which the far-field voltage pattern is given by Eq. (2). The bandwidth of the array is defined as $2\Delta f$, where Δf is the change in frequency required to frequency scan the array pattern from a maximum to the 3 db down point on the array pattern. This would require the array factor, i. e., $\cos(\pi d/\lambda \sin \theta)$, to change by $1/\sqrt{2}$. Thus

$$(7) \quad \frac{\pi d}{\Delta \lambda} \sin \theta = \frac{\pi d \Delta f}{c} \sin \theta = \frac{\pi}{4}$$

or

$$(8) \quad \text{Bandwidth} = 2\Delta f = \frac{c}{2d \sin \theta}.$$

Therefore as the distance between array elements increases and/or as the angle θ increases, the array bandwidth decreases. For $\theta = 0$, it can be seen that the array has an infinite bandwidth, assuming equal length feedlines in a corporate fed geometry, as in Fig. 1.

This is an extremely important consideration when high information rates (large instantaneous bandwidth) are required in arrays in which large baselines are also required. The insertion of phase-shifters in the lines of an array will not change the instantaneous bandwidth, but only serve to position the array pattern. However, the insertion of variable delay lines in the lines feeding an array does affect the bandwidth of the array. For example (refer to Fig. 1), to obtain an infinite bandwidth from the array at the angle θ shown, a delay line would be needed between element 1 and the summing point. The delay line would have a delay time equal to the delay time represented by the free-space path length between the plane wave front and element 2. This would effectively cause the array feed lines to be of equal length at the angle θ . Since the delay line would also affect the position of the array pattern, it would have to be very closely controlled to accomplish accurate centering of the array pattern on the signal source as the source moved. In actual practice, a stepped delay line would likely be used in conjunction with a phase-shifter, which would be used to obtain vernier control of the array pattern. With the advent of accurate, continuously controlled delay lines however, the phase-shifter might be eliminated entirely.

D. Array pattern control

There are techniques that may be used to modify the side-lobe structure of the array pattern to the extent that side-lobes are reduced considerably or eliminated effectively[13]. This is an important consideration when the array is used in a multiple target situation, where accurate position determination is needed with no ambiguities or where a single-lobed array pattern is required. To some extent the element pattern reduces the side-lobe level, but this occurs only at angles on the order of magnitude of the element half-power beam-width. For widely spaced array elements, there may possibly be numerous array lobes contained within the element half-power beam-width. Two techniques for achieving single-lobe performance from an array are given in Appendixes A and B.

CHAPTER III DESCRIPTION OF ARRAY AND DATA PROCESSING SYSTEM

A. Introduction

The subject array consists of four independently steered, 30 foot, paraboloids, the outputs of which are combined coherently at intermediate frequency with adaptive receiver circuitry. The primary function of this array is to demonstrate the feasibility of adaptive phasing techniques, and to answer some of the questions that always arise in any untried system. However, in addition to being a research tool, the system has been designed to function as an operational receiving site in a passive satellite communication system. At the present time the array has been instrumented for Echo-type reflectors, with transmitters being located at Rome, N. Y.; Trinidad, Trinidad-Tobago; and Ohio University, Athens, Ohio. In these experiments, a 60-foot parabola at Rome, N. Y., the aperture of which is equivalent to the antenna array, will function as a receiver also and hence will afford an excellent control on the performance of the array.

Expected advantages of the array are (a) ability to obtain the maximum possible gain from the array at all times, even in the presence of coherent atmospheric disturbances (phase scintillations) and of incoherent disturbances (wavefront crinkling)[14, 15] ;
(b) ability to correct for differential doppler effects occurring between

elements of the array; (c) capabilities for increasing the zone of initial signal acquisition; (d) some control over the amplitude distribution of the array aperture; and (e) the possibility of using signal-processing techniques to optimize the array for different applications, thus making a more versatile overall antenna system. Finally, the array appears to be less costly than a single large aperture, particularly for aperture sizes greater than 85 or 100 feet.

B. Overall system

1. General

The antennas are arranged on the corners of a square, sixty feet on a side, with the diagonals located in a north-south and east-west direction. The array is shown in Fig. 4. The reflectors are capable of operation at frequencies greater than 15 Gc and are steered by elevation-over-azimuth mounts. The mounts have individual servo systems but have the capability of closed-loop operation between them; in this mode one of them functions as a master and the remaining three as slaves. Input information to the master can be in the form of a remote synchro source, auto-track information, or manual position, while input information to the slaves is remote synchro source or manual position.

The antennas are focal-point fed. Equipment located at the focal point of each antenna consists of a parametric amplifier with a 2 db

noise figure, a mixer-IF preamp, test couplers, and polarization switches. In the master antenna additional equipment is present for the auto-track function. Block diagrams of the antenna feeds are shown in Fig. 5. The signals from the mixers at 30 Mc are distributed by means of coaxial cables to a common summing point located in the van at the center of the array. Local oscillator signals, as well as the various power and control signals, are fed through cables to each antenna.

Two tracking functions are required. One is that of mechanically tracking the antennas so that the element patterns will be centered on the signal source in order that maximum signal will be obtained from each antenna. The second is that of electrically tracking the interferometer lobes within the element patterns, so that one of these lobes is centered on the signal source, in order that the full gain of the antenna system as a whole is realized. In order to accomplish the electrical tracking, the 30 Mc signals are each fed to a phase-locked receiver in which the reference signals for the phase detectors are obtained from a common source. In this way each signal is locked in phase to a common source; hence, they are in phase with each other and thus may be added coherently. The mechanical tracking is accomplished by means of an amplitude-monopulse system located in one of the antennas. The data signal is tracked with the master antenna,

A. TRACKING DISH

B THREE NON-TRACKING DISHES

Fig. 5--Feed system block diagram.

and the slave antennas are kept aligned with the master by means of a common synchro transmitter in the master mount.

Acquisition in the array is accomplished by decollimating the antennas into a fan beam which is controlled in azimuth, elevation, and beam-tilt angle. In the acquisition mode the fan beam is positioned so that its wide dimension is perpendicular to the expected trajectory of the signal source. When one of the antennas acquires a signal in a threshold detection circuit associated with it, the remaining three antennas collimate with the one receiving the signal and the tracking function is initiated.

2. Monopulse tracking system

A block diagram of the monopulse tracking system is shown in Fig. 6; the equipment is located in the control van. The inputs to the tracking receiver are obtained from the master antenna feed in the form of a sum signal and two difference signals, one for elevation and one for azimuth. The sum signal acts as a phase reference and the phases of the difference signals are measured with respect to it; this gives an output to drive the antenna mount in the correct direction to keep the nulls of the difference patterns, which are made coincident with the boresight axis of the antenna, directed at the source of signal. This is accomplished by forming the complex sums as indicated in Fig. 6, which are shown in the complex plane in Fig. 7. Since the

Fig. 6--Monopulse tracking receiver.

Fig. 7--Complex vector sum used for monopulse tracking error signal.

difference signals are modulated at a 60 cps rate for the carrier servo system in the antenna mounts, the output of the phase detector, which is proportional to the cosine of the angle θ (shown in Fig. 7), is also modulated at a 60 cps rate and represents an error signal of the correct phase to drive the mount in the direction to minimize the output of the difference channel. The difference signals are either in phase with the sum signal or out of phase, depending upon which side of the difference null the signal resides, as a result of the characteristics of the split pattern in azimuth and elevation. Thus the output phase of the phase detector determines mount direction and the output magnitude determines mount velocity. The output of the phase detector is then passed through a series integrater which, together with a series integration in the servo system of the mount itself, forms a closed-loop

servo of the second type. This closed-loop servo has such characteristics that a constant error signal is required for a constant acceleration of the controlled variable; in this case, the mount motion. It can then be seen that the tracking system is a closed-loop servo system in which the loop is closed through the mount motion. The predetection bandwidth of the tracking receiver is 5 Kc and the postdetection bandwidth is on the order of 1/3 cps. This gives tracking capability on signals having a level of -160 dbm with a +10 db signal-to-noise ratio in the servo loop, when the various servo systems involved (i. e., the AFC or phase-locked local oscillator and a phase-locked receiver associated with the data output of the tracking antenna, both of which will be described later) are locked to the signal of interest[16].

3. Signal acquisition system

One of the problems encountered in the operation of high-gain antennas is that of initially acquiring the signal electrically and establishing mechanical track. With the inverse relationship between aperture size and antenna beamwidth, the larger the antenna, the more difficult is the acquisition problem. The problem is further compounded because the normal acquisition procedure for a large antenna is to scan mechanically the beam of the antenna in a circular, spiral, or zig-zag fashion about the region of space where the signal source or reflector is expected. With large antennas, this type of

scanning imposes large, mechanically undesirable stresses in the antenna mount and drive mechanism.

The method of acquisition chosen in the four-element array is to change the effective beam shape by individually pointing each of the antennas in order to increase the acquisition cone by decollimating the antennas so a fan beam is formed. Because of the uncertainty of the Echo ephemeris data, it was felt that an acquisition cone of 5° would be needed for consistent acquisition. Manual synchros provide the initial azimuth and elevation look angles as obtained from ephemeris data. A resolver is used to divide a reference voltage into sine and cosine components so that the tilt of a fan beam can be controlled, and offset networks are used for offsetting the antenna axes with respect to each other so that a fan beam is formed. In the azimuth channel, a secant correction is introduced to control the azimuth tracking loop gain as a function of the secant of the elevation angle, and also to control the amount of azimuth offset as a function of elevation angle. A block diagram of the acquisition equipment is shown in Fig. 8. The geometry of the fan beam is shown in Fig. 9. The inclination of the fan beam is set so that maximum intercept area is afforded by the four beams, that is, the narrow dimension of the fan beam is set to be coincident with the satellite trajectory. When one of the four antennas receives a signal, its threshold detector, shown as an out-rigger to each

Fig. 8--Acquisition equipment.

receiver in Fig. 10, controls a respective relay in the acquisition equipment and causes two events to occur. First the entire fan beam is caused to scan in the direction of the satellite trajectory with a set-in initial angular velocity. Simultaneously, the three antennas which did not receive signals collimate with the antenna that did receive the signal.

Fig. 9--Fan beam geometry.

Simultaneously with the controlling of the antenna position, the signal acquisition equipment controls the local oscillator. As will be described in the signal cohering section, the local oscillator is swept at a 1 cps rate in order to acquire the doppler signal. When one of the slave antennas receives a signal, the sweep is stopped at the frequency where the signal was first detected. As the antennas converge or collimate and the tracking antenna receives the signal, the

Fig. 10--Signal cohering and combining receiver.

threshold detector in that receiver engages a relay which transfers the control of the VCO in the local oscillator to either the AFC discriminator or the phase detector in the signal processing equipment, depending upon which mode is selected. In addition to the above method of acquisition, an optical telescope which drives synchro transmitters, is available for acquisition and visual tracking when ephemeris data are not available, or as a backup to the electronic acquisition equipment.

4. Signal-tracking and cohering equipment

In order to cohere the signals from the four antennas, each signal must be at the same frequency so that phase will have some significance. Two factors must be considered when this requirement is imposed. The first is that of coherent doppler shift (± 100 Kc for the planned Echo II experiments), which must be corrected for in order that the data receivers may have a minimum pass band. The second factor is that of differential doppler shifts between elements of the array, caused by the different path lengths involved in the satellite-antenna geometry[17]. This differential doppler can reach a maximum of 2 cps in the Echo II experiments.

The coherent doppler shifts appearing at each antenna element are corrected by a shift in the frequency of the first local oscillator shown in Fig. 11. The local oscillator itself is frequency- and phase-locked to a crystal oscillator in order to impart to the output of the power tube

Fig. 11--Local oscillator.

the stability that is inherent in the crystal oscillator. In the internal servo loop of the local oscillator, the gain of the AFC loop is made low and the gain of the phase-locked loop is made high in order to assist in the tuning of the power oscillator to the correct frequency and to aid in the maintenance of phase-lock for large frequency errors by extending the lock-on range of the loop up to 100 Kc on either side of the correct frequency. Doppler shifts of up to ± 100 Kc are expected to occur in the Echo II experiments.

A voltage-controlled oscillator in the local oscillator, which controls its output frequency, may be controlled from any one of three different sources, depending upon the mode of operation desired. In the acquisition mode a 1 cps oscillator having a triangular output waveform is used to drive the control voltage-controlled oscillator (VCO). This causes the local oscillator to sweep in frequency up to ± 100 Kc about the center operating frequency in order to capture the satellite signal in the presence of doppler shift. When the signal is received, the search stops and the VCO input is switched to a discriminator or a phase detector, or both, in the data receiver, depending upon which mode is desired. The sweep range and rate of the VCO can be controlled to optimize the acquisition under various conditions.

The signal-cohering equipment is shown in block diagram form in Fig. 11. Three complete phase-lock loops and one partial loop are

shown. The partial one is shown in Channel A, the tracking antenna; the remaining section of the loop is the local oscillator. The loop is closed through either the phase detector or the discriminator shown in Channel A. The 28 Mc crystal oscillator is not presently used as a voltage-controlled oscillator but can be used as one with the loop then closed through the phase detector in that channel. The remaining three phase-lock loops are located in the slave channels, one in each channel, in order to correct for vernier phase (or frequency) effects that may occur between these antennas and the tracking antenna.

The reference for the phase detectors in the slave antennas may be obtained in three possible ways. In mode I, the reference signal is the master IF signal so that each antenna is locked to the tracking antenna signal. In mode II, each antenna signal is locked to the sum of the four IF signals[18], and in mode III, each antenna is locked to a stable crystal oscillator. In modes I and II the local oscillator can be frequency-locked to the tracking antenna signal, phase-locked to it, or both, in any mode. These three modes of operation allow considerable flexibility in the system when used as a research tool and allow the array performance to be optimized under different operational conditions. For example, in mode I, with the addition of a narrow-band filter, the four receivers may be used on incoherent signals, such as those encountered in radio astronomy, thus giving the potential of a 60-foot antenna for radio-astronomy use.

As can be seen, the signals from the four receivers are combined at the IF frequency and then detected. This arrangement permits maximum signal-to-noise ratio improvement. When the signals are summed, the array pattern takes the form of several interferometer lobes which have as an envelope the response of the single elements (Fig. 12). As

Fig. 12--Element and array antenna pattern.

pointed out previously, the positions of the element patterns are controlled by the mount positions, while the locations of the interferometer lobes are controlled by the signal-cohering equipment and are adaptively scanned electronically by this equipment.

As an alternative to summing the signals at IF before detection, it is possible to sum them after detection[19, 20], or to multiply them before or after detection. Either synchronous detectors, such as the phase-lock demodulators shown in Fig. 13, or conventional envelope

Fig. 13--Phase locked demodulator with quadrature detector.

detectors can be used. The various available ways of combining the signals make it possible to alter the pattern response of the array in a flexible way and hence to optimize the array for different applications, such as for tropospheric scatter, multiple-target situations, position determination, detection of coherent or incoherent point sources embedded in noise backgrounds, etc. This extensive capability for processing the signals from the elements of the array puts the antenna system into the class of signal-processing antennas[21].

The complete electronics system for the array, including the acquisition equipment, antenna-mount controls, signal-processing equipment, local oscillator, and tracking circuits are shown from left to right, respectively, in Fig. 14. The data-recording equipment is shown in Fig. 15. The statistical analyzing system is shown in Fig. 16. This is used to analyze the data which are recorded on magnetic tape. The system will provide narrow and broadband probability density functions, power and cross-power spectral densities, and auto- and cross-correlation functions.

Fig. 14--Overall view of mount control,
cohering receiver and local oscillator.

Fig. 15--Overall view of
recording equipment.

Fig. 16--Overall view of data analysis system.

CHAPTER IV

DETAILED SYSTEM ANALYSIS

A. Introduction

From the preceding chapter, which described in some detail how an adaptively phased array might be instrumented, it is clear that free use has been made of the phase-locked loop. In fact, several different uses of it have been employed to demonstrate its usefulness in a variety of situations. These uses in the array might be outlined as follows:

- (a) demodulation
- (b) phase control of signals with respect to a reference, and
- (c) synchronization and stabilization of oscillators.

In these three general categories, the loops have been used to provide the functions of a narrow-band tracking filter; as detectors of frequency, phase, and amplitude modulation; automatic frequency control; and frequency translation. As a result of this tremendous versatility in a variety of uses, the phase-locked loop has come to be referred to by several terms, among which are phase-coherent loop, active filter, dynamic filter, tracking filter, prediction filter, correlation detector, coherent detector, synchronous detector, product detector, keyed detector, and commutation detector. The above uses and terminology reflect three concepts; i. e.,

- (a) a locking on or synchronization,
- (b) a dynamic aspect to their operation, and
- (c) a correlation of the present with the past
and a prediction of the future.

The purpose of this chapter is to analyze the operation of the phase-locked loop (PLL) and to show the interrelationship among the above three concepts and how they might be controlled.

B. Linearization

A typical PLL is shown in Fig. 13, with an associated quadrature detector. The basic operation of the loop is to maintain a fixed phase relationship between the input signal and the voltage-controlled oscillator; this is accomplished by the action of the feedback loop. The loop and the effects of the feedback can easily be analyzed by standard control system techniques, particularly by means of the root locus and the Bode plot. However, before these techniques can be applied the loop must be linearized, since the phase detector is nonlinear in its operation. This has been done by Weaver[22] and by Jaffe and Rechtin[23]. The technique used in this paper will follow that of Weaver.

Let the input signal to the loop be represented by a carrier-plus-noise given in Eq. (9),

$$(9) \quad E_{in} = \sqrt{2} E_s \sin (\omega_s t + \theta_1) + N(t) ,$$

and the oscillator signal by Eq. (10),

$$(10) \quad E_O = \sqrt{2} E_O \cos (\omega_{st} + \theta_2) \quad .$$

The incoming signal-plus-noise is multiplied in the phase detector by the output of the VCO. The result for signal alone is

$$(11) \quad E_{p_s} = K_m E_O E_s [\sin (\theta_1 - \theta_2) + \sin (2\omega_{st} + \theta_1 + \theta_2)] \quad .$$

Since the filter in the loop is of the low-pass type, it is designed to suppress the second harmonic term with the result that

$$(12) \quad E_{p_s} = K_m E_O E_s \sin (\theta_1 - \theta_2) \simeq K_m E_O E_s (\theta_1 - \theta_2)$$

for small $(\theta_1 - \theta_2)$. The constant K_m is the phase detector constant.

The component due to noise may be found by expanding $N(t)$ in a complex Fourier series:

$$(13) \quad N(t) = \sum_K C_K e^{i(\omega_K + \omega_s)t} \quad .$$

The output of the phase-detector due to noise can then be found as

$$(14) \quad E_{p_N} = K_m E_O \frac{N'(t)}{\sqrt{2}} \quad .$$

The sum of the low-frequency components is then

$$(15) \quad K_m E_O \left[E_s (\theta_1 - \theta_2) + \frac{N'(t)}{\sqrt{2}} \right] \quad .$$

If there were modulation on the input signal, this would have to be included in Eq. (9). The phase detector may now be converted to its linear equivalent, a summing junction, by setting the input to the loop equal to

$$(16) \quad \theta_1 + \frac{N'(t)}{E_s \sqrt{2}} ,$$

where $N'(t)$ is the narrow band $N(t)$ translated down to zero frequency with a phase change equal to θ_2 ; and by adding inside the loop an amplifier having a gain given by Eq. (17):

$$(17) \quad K_3 = K_m E_o E_s .$$

The VCO changes frequency with a change in its input voltage, which causes its transfer function to be that of an integrator between the input voltage and the output phase. Hence the PLL can be represented as in Fig. 17(a) with all gain constants separated and as in Fig. 17(b) with all gain constants grouped into two amplifiers. It should be noted at this point that the loop gain is dependent upon the input signal strength and it is this feature which causes the loop to have adaptive properties. The closed-loop transfer function is that between C and R, and the open-loop transfer function is that between M and E with the loop opened at the negative input to the summing port.

(a)

$$K = K_m E_s E_0 K_1 K_2 = K_a K_b$$

(b)

Fig. 17--Linearization of the phase locked loop.

C. Root locus method of analysis[24, 25]

As can be seen by examination of Fig. 17(b), the two parameters in the loop which might be varied are the loop gain and the transfer function of the filter. The main object in the design of a loop is to achieve satisfactory transient response as well as to control the magnitude of the error signal to stay within the linear range of the phase detector.

The closed-loop transfer function of the loop in Fig. 17(b) is given by

$$(18) \quad \frac{C(s)}{R(s)} = \frac{K_a G(s)}{1 + K G(s) H(s)},$$

where

$$(19) \quad K_a G(s) = K_a \frac{\prod_j (s + a_j)}{\prod_K (s + b_K)} = K_a \frac{Z_g}{P_g}$$

and

$$(20) \quad K_b H(s) = K_b \frac{\prod_i (s + \alpha_i)}{\prod_l (s + \rho_l)} = K_b \frac{Z_h}{P_h}$$

and where the Z's denote the product of zero factors and the P's denote the product of pole factors [26]. By substituting Eqs. (19) and (20) into Eq. (18) and reducing, there can be obtained

$$(21) \quad \frac{C(s)}{R(s)} = K_a \frac{Z_g P_h}{P_g P_h + K Z_g Z_h},$$

demonstrating that the zeros of the feed forward path and the poles of the feedback path become zeros of the closed-loop transfer function.

To find the poles of C/R , the denominator of Eq. (21) is set equal to zero, which is equivalent to

$$(22) \quad K G(s) H(s) = -1 = 1 \angle 180^\circ,$$

where $K G(s) H(s)$ is the open-loop transfer function. The root locus is the locus of closed poles in the complex plane as the gain K varies. Hence the loci are the points in the S -plane where the angle of $G(s) H(s)$ is 180° .

To find the S -plane phase angle, the angles from the point of interest to the zeros of $G(s) H(s)$ are added and, from this sum, the sum of the angles from the point of interest to the poles of $G(s) H(s)$ are subtracted. The angles are measured in a counter-clockwise direction from a line parallel with the real axis. These angles can easily and quickly be measured with a Spirule[27] , a patented device manufactured by a company of the same name. Figure 18 demonstrates

Fig. 18--Measurement of angles for the root-locus.

the measurement of the above angles, where the angle ψ must satisfy the relationship

$$(23) \quad \psi = \theta_1 - \phi_1 - \phi_2 - \phi_3 = 180^\circ$$

for the point S_1 to lie on the root locus. Although it may appear at first that the root locus would be difficult to obtain, a few simply applied rules[28] are of considerable aid in sketching in the locus.

These might be listed as:

- (a) The number of branches of the root locus is equal to the number of closed-loop poles, where a branch is a separate portion of the root locus which represent all values of K from 0 to ∞ .
- (b) Each branch of the root locus starts at an open-loop pole at $K = 0$ and ends at an open-loop zero or infinity with $K = \infty$.
- (c) For a locus to exist on the real axis, the sum of the poles and zeros to the right of the point on the axis must be odd.
- (d) The root locus is symmetrical with respect to the real axis.

(e) The angle by which the root locus leaves an open-loop pole or approaches an open-loop zero is in the direction $180^\circ q$, minus the sum of the angles of the vectors from the remaining poles and zeros to the pole or zero in question, where q is an odd integer.

(f) The point at which the root locus leaves the real axis is determined by equating the reciprocals of the distances from the poles and zeros on the real axis to zero.

(g) The direction of the asymptote lines to the root locus is given by $\pm 180^\circ q / (n - m)$, where n is the number of open-loop poles, m is the number of open-loop zeros and q is an odd integer.

(h) The asymptote lines cross the real axis at the point determined by the relationship (sum of real parts of the poles - sum of real parts of the zeros) / (total number of poles - total number of zeros).

A typical root locus plot is shown in Fig. 19, with the position of the poles and zeros corresponding to those of the phase-locked loop shown in Fig. 17(b). The pole at the origin is due to the VCO and the pole and zero on the negative real axis correspond to that of a lag network which is frequently used as the feed forward filter in the loop.

Fig. 19--Typical root-locus for phase locked demodulator.

When the root locus has been plotted, as in Fig. 19, the gain required to move a closed-loop pole to a specified position may be obtained by fulfilling the second condition of Eq. (22), that is,

$$(24) \quad |G(s) H(s)| = \frac{1}{K} \quad .$$

This can be easily accomplished by referring to Fig. 18, where the function, whose poles and zeros are shown in the figure, can be represented as

$$(25) \quad F(s) = \frac{(s - s_2)}{(s - s_3)(s - s_4)(s - s_1)(s - s_5)} \quad .$$

There will be a corresponding term in the time function of the form

$$(26) \quad f_1(t) = \frac{s_1 - s_2}{(s_1 - s_3)(s_1 - s_4)(s_1 - s_5)} e^{-s_1 t},$$

where attention is directed at the behavior of the function at $s = s_1$.

In general, s_1, s_2 , etc. will be complex numbers, so Eq. (26) can be written as

$$(27) \quad f_1(t) = \frac{|s_1 - s_2|}{|s_1 - s_3| |s_1 - s_4| |s_1 - s_5|} e^{-s_1 t + \psi},$$

where ψ is given by Eq. (23). By referring to Fig. 18, it can be seen that $|s_1 - s_2|$, $|s_1 - s_3|$, etc. are the lengths of the vectors from s_1 to s_2, s_3 , etc., respectively. Thus the magnitude of a complex function such as Eqs. (24) or (25) is the product of the vectors from the point of interest to the zeros divided by the product of the vectors from the point of interest to the poles [29]. The points of interest in the root loci are those points which comprise the loci of the closed-loop poles, i. e., the root loci themselves.

Thus vector multiplication and division can be quickly and accurately performed by use of the Spirule. Expressed analytically, the loop gain required to move a closed-loop pole to a specified point on the root locus is given by

$$(28) \quad K = \frac{\prod d_K}{\prod l_j}$$

where d_K is the distance from the point to the poles and l_j is the distance from the point to the zeros.

The root locus plot indicates many of the network's transient characteristics of the network. For example, the real component of the closed-loop poles gives directly the exponential decrement rate, while the imaginary part gives the natural frequency. Thus, for a pole at $s = -\alpha + j\beta$, the transient term corresponding to this pole would be $A e^{(-\alpha + j\beta)t}$. From Fig. 20 it can be seen that lines of

Fig. 20--Definition of damping factor, undamped natural frequency, damped natural frequency, and decrement factor.

constant decay rate are vertical and the farther from the imaginary axis, the more rapid the decay. The value of the natural frequency establishes the sinusoidal oscillation frequency of the system; this is given by horizontal lines, on which the frequency is constant, and the greater the distance from the real axis, the higher the oscillation frequency. Also from the root locus plot can be obtained the decrement factor ζ and the undamped natural frequency where

$$(29) \quad \zeta = \frac{\text{actual damping rate}}{\text{undamped natural frequency}} = \frac{\alpha}{\omega_0}$$

and

$$(30) \quad \omega_0 = \sqrt{\alpha^2 + \beta^2} \quad .$$

The values of ζ and ω_0 are customarily associated with the characteristic equation of a second-order system:

$$(31) \quad s^2 + 2 \zeta \omega_0 s + \omega_0^2 = 0 \quad .$$

From the root locus, then, the decay rate, α ; the decrement factor, ζ ; the undamped natural frequency, ω_0 ; and the natural frequency, β , can be found. From these parameters, the response of the system can be obtained for a step function input. From Fig. 21, it can be shown that [30]

$$(32) \quad t_p = \frac{\pi}{\beta} \quad , \quad t_s = \frac{3.0}{\zeta \omega_0}$$

Fig. 21--Step function response measurement.

where t_p is the time from the application of the step function to the first peak and t_s is the time from the application of the step to the point at which the response has decayed to 5%. The overshoot as a function of the decrement factor is given in Table I.

TABLE I
THE DECREMENT FACTOR ζ VERSUS THE BANDWIDTH
AND THE PEAK OVERSHOOT RATIO (POR)

Decrement factor ζ	ratio of bandwidth ω_b to ω_o	Maximum Overshoot	Number of oscillations to settle to 5%
0	1.55	1.0	∞
0.1	1.54	0.730	4.76
0.2	1.51	0.527	2.34
0.4	1.39	0.254	1.09
0.5	1.27	0.166	0.82
0.6	1.13	0.0955	0.64
0.707	1.00	0.042	0.48
0.8	0.90	0.016	0.36
1.0	0.2	0.00	0.00

D. Bode plot

The Bode plot[31] furnishes a convenient way of viewing the operation of a control system in the frequency domain. It consists merely of a plot of the magnitude and phase of a transfer function as a function of frequency. For the open-loop transfer function, the Bode plot can be easily drawn by noting the various time constants in the loop and assigning a break-frequency to these time constants. The Bode plot can then be sketched, making it asymptotic with the break-frequency lines. The construction of the Bode plot will not be described in detail since it is a rather commonly understood procedure.

The closed-loop Bode plot is more difficult to obtain, since the closed-loop poles are generally unknown initially. Use can be made of Nichols charts for transferring between open-loop and closed-loop Bode plots, but it is felt by the author that perhaps an easier way would be to obtain the values of the closed-loop poles from the root loci and the closed-loop zeros from inspection and then construct the closed-loop Bode plot by conventional asymptotic techniques.

The Bode plot can also be easily obtained from the root locus by determining the magnitude and phase of the pole zero configuration along the $j\omega$ axis[32] . When this is done, the feedback poles become closed-loop zeros, the open-loop zeros stay as closed-loop zeros, and the closed-loop poles are found on the root locus corresponding to open-loop gain. The determination of magnitude and/or phase can again be quickly and accurately accomplished by use of the Spirule.

The Bode plot of a phase-locked loop can be easily measured by phase or frequency modulating the input signal with a variable frequency ω_0 . By noting the ratio of the output to input modulation amplitude as a function of ω_0 , the amplitude portion of the Bode plot may be obtained; and by noting the phase shift as a function of frequency, the phase portion of the Bode plot may be obtained. From the amplitude portion of the Bode plot, the location of the open-loop zeros and closed-loop poles can be found. By constructing the root locus

from these, the open-loop poles can be found. Once this is known, the operation of the loop can be easily determined and changed to suit design considerations.

The Bode plot is a particularly convenient way to visualize the operation of a PLL when functioning as a filter or when a particular frequency characteristic such as bandwidth is desired. The root locus plot is convenient for obtaining the closed-loop poles from the open-loop transfer function, for obtaining the transient response of the loop, for visualizing the effect of a change in signal strength and/or loop gain, and for quickly determining the tracking range of a PLL. In the analysis to follow, both techniques will be used. The purposes of the following sections are to give insight into the design of PLL for various purposes, to point out the various important considerations, and to show how these are influenced by the loop variables, i. e., gain and filter response. It has been the experience of the author and of others[33] who have worked with such loops, that the best design is one of cut and try guided by a knowledge of the various effects that can be controlled by changing the loop parameters.

Perhaps the reason for this is the fact that many effects are extremely difficult to include in an analytical treatment of the loops. Some of these effects might be caused by nonlinearities, stray capacitance, loading effects caused by cascading components,

saturation, cross-coupling among components, and less than theoretical performance from the loop components.

E. Transient response

One of the design goals in any closed-loop control system is to provide adequate transient response. This situation is true also for the phase-locked loop, where the transient response is one of two major design objectives. The second objective is to provide adequate error signal characteristics. This section will detail how the root locus can be used to design and/or calculate transient response. For example, take the phase locked demodulator as shown in Fig. 13 and assume the feed forward filter to have the transfer function

$$(33) \quad G(s) = \frac{s + a}{s + b},$$

where $b < a$. The closed-loop transfer function is then

$$(34) \quad \frac{C(s)}{R(s)} = \frac{K_a s(s + a)}{s(s + b) + K(s + a)},$$

and the root locus for the closed-loop poles is shown in Fig. 19.

Assume further that the input signal is changing in frequency steps, such as would be encountered in frequency-shift carrier modulation. This would cause the Laplace transform of the input frequency to be of the form

$$(35) \quad \omega(s) = \frac{A}{s},$$

and, since the phase is related to the frequency by

$$(36) \quad \phi(s) = \frac{\omega(s)}{s}$$

the Laplace transform of the input phase will be given by

$$(37) \quad R(s) = \phi(s) = \frac{A}{s^2}$$

Then the output $C(s)$ can be obtained from Eq. (34) as

$$(38) \quad C(s) = \frac{A}{s} \frac{K_a(s+a)}{s(s+b) + K(s+a)}$$

From this equation it can be seen that the transient response $C(t)$ will be that of a network having a transfer function $\frac{K_a(s+a)}{s(s+b) + K(s+a)}$ excited by a step function.

Depending on the value of the open-loop gain, the closed-loop poles are found from Fig. 19, for example, at $-\alpha \pm j\beta$. There is a closed-loop zero at $s = -a$, and a third closed-loop pole at $s = 0$ is caused by the step function excitation. The transient response can be calculated from the factored form of Eq. (38), which is

$$(39) \quad C(s) = \frac{K_a A (s+a)}{s(s+\alpha+j\beta)(s+\alpha-j\beta)}$$

The transient response is then[34]

$$(40) \quad C(t) = C_1 + C_2 e^{-\alpha} e^{-j\beta t} + C_3 e^{-\alpha} e^{+j\beta t}$$

with

$$C_1 = \frac{K_a A (s + a)}{(s + \alpha + j\beta)(s + \alpha - j\beta)} \bigg|_{s=0},$$

$$C_2 = \frac{K_a A (s + a)}{s(s + \alpha - j\beta)} \bigg|_{s = -\alpha - j\beta}, \text{ and}$$

$$C_3 = \frac{K_a A (s + a)}{s(s + \alpha + j\beta)} \bigg|_{s = -\alpha + j\beta},$$

and where the C 's are merely the residues of Eq. (39) at the poles. The C 's may also be determined as in Eq. (27), or by the use of the Spirule, by finding the product of the distances from the zeros to the pole of interest and dividing by the product of the distances from the remaining poles to the pole of interest.

The peak over-shoot ratio, time to reach peak, and settling time are related to the position of the complex closed-loop poles, as shown in Table I and Eq. (32). Thus, one consideration in choosing the open-loop gain is the location of the closed-loop poles on the root locus to give desired peak overshoot, time to peak, and settling time.

As a second example, the transient response of the same demodulator will be studied with a frequency ramp input, i. e.,

$$(41) \quad \omega(s) = \frac{A}{s^2} \quad .$$

The corresponding phase transform is

$$(42) \quad R(s) = \phi(s) = \frac{A}{s^3} \quad ,$$

which gives, for $C(s)$,

$$(43) \quad C(s) = \frac{A}{s^3} \frac{K_a s(s+a)}{s(s+b) + K(s+a)} \quad .$$

This is equivalent to exciting a network with the transfer function

$$\frac{K_a (s+a)}{s[s(s+b) + K(s+a)]} \quad \text{with a step input.}$$

The values of the closed-loop poles for the bracketed term are found from the root locus of Fig. 19. The value of $C(s)$ is then found to be

$$(44) \quad C(s) = \frac{A K_a (s+a)}{s^2 (s + \alpha + j\beta) (s + \alpha - j\beta)} \quad .$$

The transient response is given by the inverse Laplace transform,

$$(45) \quad C(t) = C_{11}t + C_{12} + C_1 e^{-\alpha - j\beta} + C_2 e^{-\alpha + j\beta},$$

$$C_{11} = \frac{A K_a (s + a)}{(s + \alpha + j\beta)(s + \alpha - j\beta)} \Big|_{s=0},$$

$$C_{12} = \left[\frac{d}{ds} \frac{A K_a (s + a)}{(s + \alpha + j\beta)(s + \alpha - j\beta)} \right] \Big|_{s=0},$$

$$C_1 = \frac{A K_a (s + a)}{s^2 (s + \alpha + j\beta)} \Big|_{s = -\alpha + j\beta}, \quad \text{and}$$

$$C_2 = \frac{A K_a (s + a)}{s^2 (s + \alpha - j\beta)} \Big|_{s = -\alpha - j\beta},$$

and where again the residues can be found by direct measurement or by use of the Spirule. As before, one of the considerations in choosing the open-loop gain K would be the transient response characteristics of the network.

F. Error signal considerations

As noted previously, one of the two major design goals in the PLL is to design for proper error signal characteristics [35]. This is of particular importance since the analysis is based on a linearized model of the PLL, with the validity based on error signal characteristics. The output of the phase detector or multiplier in the loop follows a sinusoidal response, i. e., $K_m E_s E_o \sin(\theta_1 - \theta_2)$, as in

Eq. (12). For the linear approximation to apply with accuracy, the range of $\theta_1 - \theta_2$ should be restricted to a value of approximately $\pm \pi/4$. This is also important when low distortion is a requirement when the loop is used as a demodulator.

The transfer function between the error signal and the reference input can be expressed as

$$(46) \quad \frac{E(s)}{R(s)} = \frac{1}{1 + K G(s) H(s)} .$$

Assume that $G(s) = s+a/s+b$, where $b < a$, and the configuration is the loop used as a demodulator. Then

$$(47) \quad \frac{E(s)}{R(s)} = \frac{s(s+b)}{s(s+b) + K(s+a)} .$$

The root locus for the closed-loop poles is again given by Fig. 19 and is identical with that for $C(s)/R(s)$ for the same configuration.

Assume again the two cases considered previously for the input, i. e., a frequency step and a frequency ramp. For the frequency step, $R(s) = A/s^2$, which gives for the error signal transform

$$(48) \quad E(s) = \frac{A}{s} \frac{(s+b)}{s(s+b) + K(s+a)} .$$

The final value theorem can be used to find the steady-state error

$$(49) \quad e_{ss}(t) = \lim_{s \rightarrow 0} s E(s) = A \frac{s+b}{s(s+b) + K(s+a)} \bigg|_{s=0} = \frac{A b}{K a} .$$

From Eq. (28), the open-loop gain K can be expressed as

$$(50) \quad K = \frac{\prod_j^d K}{\prod_j l_j} = \frac{r_1 r_2}{r_3},$$

where r_1 , r_2 , and r_3 are defined in Fig. 19. Making the substitution of Eq. (50) into Eq. (49),

$$(51) \quad e_{ss}(t) = \frac{A b r_3}{r_1 r_2 a},$$

where it is required, by proper pole and zero position and open-loop gain control, to restrict this quantity to less than $\pi/4$. One obvious way to accomplish this would be to make the transfer function of the feedforward filter to be that of an integrator, i. e., $G(s) = s+a/s$, which would make $b = 0$ and likewise reduce the steady-state error to zero. If this is not done, it can be easily seen that moving the pole position of $G(s)$ far to the right and the zero position of $G(s)$ far to the left on the $-\sigma$ axis will tend to reduce the error.

Consider the error signal characteristics with a frequency ramp input. This is equivalent to the loop tracking a doppler signal and is an important input in the design of PLL for use with high-velocity satellites. The phase input for a frequency ramp is $R(s) = A/s^3$, the error signal is of the form

$$(52) \quad E(s) = \frac{A}{s^3} \frac{s(s+b)}{s(s+b) + K(s+a)},$$

and the steady state error is

$$(53) \quad e_{ss}(t) = \lim_{s \rightarrow 0} s E(s) = \frac{A(s+b)}{s[s(s+b) + K(s+a)]} \bigg|_{s=0}.$$

In this situation, the feedforward transfer function $G(s)$ is required to be that of an integrator for any control of the error signal, i.e., b must equal zero. Thus

$$(54) \quad e_{ss}(t) = \frac{A}{K a} = \frac{A}{r_1^2}.$$

Control of the error signal is therefore dependent on the open-loop gain, which determines the value of r_1 for a fixed zero position.

The length of r_1 can be viewed as proportional to the bandwidth of the loop, where an increased bandwidth will reduce the steady-state error signal as well as admit more noise. Hence there is a compromise in the choice. Many times it is found that the bandwidth of the loop required to give adequate transient response is less, by a considerable amount, than that required to give adequate error signal characteristics.

G. Effect of signal strength on bandwidth

As has been noted, the loop bandwidth determines the transient characteristics and error signal bounds. Also it determines the output signal-to-noise ratio, and, as such, it is important to appreciate how this changes. Since the root locus is merely a plot of the locus of closed-loop poles, it is difficult to derive quickly the bandwidth of a

pole-zero configuration. However it can be said on a qualitative basis that the distance from the origin of the complex plane to the closed-loop root is proportional to the bandwidth. Quantitative information on the bandwidth, however, can quickly be obtained from the Bode plot, which is accurately plotted by the use of a Spirule moving along the $j\omega$ axis. Consider again the closed-loop transfer function of the PLL as given in Eq. (34), where the transfer function is in terms of the phase of the input. The transfer function for a frequency input is given by

$$(55) \quad \frac{C(s)}{R(s)} = \frac{K_a(s + a)}{s(s + b) + K(s + a)} .$$

The closed loop poles are found on the root locus of Fig. 19 according to the open-loop gain K , and the open-loop zero remains as a closed-loop zero. Figure 19 is replotted in Fig. 22 with specific values of the gain K indicated on the locus. The corresponding Bode plot is shown in Fig. 23 as a function of the open-loop gain K . The bandwidth of the closed-loop Bode plot is defined as the frequency at which the response is 3 db down from that at zero frequency[36]. As can be clearly seen, the bandwidth increases with the open-loop gain, however the relationship is not linear. If the system were a second-order system with no zeros, the bandwidth would be proportional to the radial distance from the origin of the complex plane to the dominant

Fig. 22--Root-locus with open loop gain.

complex poles as given by Table I. However, as can be seen, the presence of a zero in the closed-loop transfer function alters this relationship.

The Bode plot in Fig. 23 gives the filter pass band of the PLL. As can be seen, the pass band varies with the open-loop gain, which is directly proportional to the input signal strength. Thus, for example, from Fig. 23, an increase in signal strength of 3 (5 db) will more than double the loop bandwidth. In this sense the loop is adaptive in that for strong signals the loop bandwidth is large,

permitting fast transient response; the error signal remains within the linear range or, alternatively, permits an increase in the rate of information transfer and keeps the transient and error response within limits. Likewise, for a weak signal, the bandwidth becomes more narrow, thus permitting less noise through the PLL. For quantitative results, both the root locus and the Bode plot should be consulted.

Since the filter pass band is given by the closed-loop Bode plot, some information as to the synthesis of a particular filter characteristic can be obtained from the root locus and Bode plot. For example, if it were desired to suppress one frequency and to pass a lower frequency, the closed-loop zero should be placed nearer the origin on the root locus; the radial distance from the complex plane origin to the closed-loop poles should be made small, but yet should be larger than the distance from the origin to the zero. This is illustrated in Fig. 23, where, for $K = 1$, the closed-loop complex poles are nearer the origin than the zero. The asymptotes for $K = 1$ are shown as dashed lines. The break upward due to the zero occurs after the break downward due to the complex poles and hence reduces the amount of attenuation in the high-frequency region. By adding additional closed-loop poles and zeros to the closed-loop transfer function, the filter function can be shaped to fit a particular requirement. The closed-loop Bode plot of the PLL to phase modulation is shown in Fig. 24.

Fig. 23--Bode plot of phase locked demodulator with frequency input .

Fig. 24--Bode plot of phase locked demodulator with phase input.

In many cases, it is desirable to keep the closed-loop bandwidth constant thereby maintaining transient characteristics and error signal response constant as a function of signal strength. When this is desired, the incorporation of an AGC control preceding the loop will do much to maintain a constant operating point. By suitable filtering, the AGC response can be made to respond to signal strength only. It is interesting to note that the phase-locked loop with a quadrature detector will permit the measurement of signal levels and, in fact, perhaps the major use of the quadrature detector is to provide for AGC control. Use can also be made of a limiter at the input to the PLL to provide for a constant input power, thus tending to keep the PLL at a constant operating point[37] .

H. Tracking range

Since one of the principal functions of the PLL is to track a signal which is varying in frequency and/or phase, it is desirable to investigate the loop from the aspect of determining the frequency range over which the loop will track and how this might be controlled[38] . The loop tracks a signal by virtue of the voltage-controlled oscillator changing its frequency and hence its phase. Thus the range over which the loop will track is dependent upon the magnitude of the voltage used to control the VCO and this is dependent upon the open-loop gain K .

According to Eq. (24) and Fig. 19, the condition for a closed-loop pole to be located at P_0 is

$$(56) \quad \left| G(s) H(s) \right|_{s=P_0} = \frac{1}{K_m E_0 E_s K_1 K_2} = \frac{r_3}{r_1 r_2} .$$

The d-c gain of the feedforward filter $G(s) = s+a/s+b$ is unity, hence

$$(57) \quad \lim_{s \rightarrow 0} K_1 G(s) = \lim_{s \rightarrow 0} K_1 \frac{s+a}{s+b} = K_1 \frac{a}{b} = 1$$

or

$$(58) \quad K_1 = \frac{b}{a} .$$

Since the maximum voltage from the phase detector is $\pm K_m E_0 E_s$, and since the feedforward filter has unity d-c gain, the maximum frequency shift of the VCO is $\pm K_m E_0 E_s K_2$ and the maximum tracking range is

$$(59) \quad f_t = 2 K_m E_0 E_s K_2 .$$

From Eqs. (56), (58), and (59), it can be shown that

$$(60) \quad f_t = 2 r_1 \frac{r_2 / b}{r_3 / a} .$$

Thus the tracking range is directly proportional to r_1 , r_2 , and a , and is inversely proportional to r_3 and b (these parameters are defined in Fig. 19). Equation (60) is valid, provided, of course, that the VCO can be physically pulled in frequency by the amount indicated

in the equation. Thus, although Eq. (60) indicates that by making the feedforward transfer function $G(s)$ an integrator, i. e., $b = 0$, the tracking range could be made infinite, the physical characteristics of the VCO would surely limit this to the range over which it would continue to act as an integrator between input voltage and output phase. In practice, the tracking range, f_t , should be made much larger than that required to accomplish doppler tracking or demodulation.

It can be shown that the open-loop gain K changes instantaneously with a change in the input phase. The output of the multiplier is $K_m E_o E_s \sin (\theta_1 - \theta_2)$. For a fixed θ_2 ,

$$(61) \quad K = \frac{\Delta e_o}{\Delta \theta_1} E_o E_s K_1 K_2 = K_m E_o E_s K_1 K_2 \cos (\theta_1 - \theta_2).$$

From this it can be seen that as θ_1 increases, the loop gain decreases on an instantaneous basis, and the root locus should be consulted for the quantitative effects produced.

I. Noise effects

The phase-locked loop is said to be above the threshold or "in-lock" when the error signal is less than a specified amount. When the phase error exceeds $\pi/2$ radians, the sign of the open-loop gain changes, thereby causing the loop to become unstable. It is commonly specified that the phase error be less than $\pi/2$ radians; but when this is done, there is an appreciable nonlinearity

present because the $\sin (\theta_1 - \theta_2)$ term represents the phase detector.

This nonlinearity will partially void the linearized model of the PLL and will also introduce distortion when the loop is utilized as a demodulator. Hence it is felt that requiring the phase error to be less than $\pi / 4$ radians will minimize the above two problems.

Two effects work together to produce the phase error: modulation and noise. Since modulation-induced phase errors are readily handled with deterministic type signals, they will not be discussed further. However to obtain a measure of the effects produced by noise, a statistical approach must be used. It has been shown by Bennett [39] that narrow-band noise, which is normally the type encountered at the output of an IF strip which precedes a PLL, can be represented as

$$(62) \quad n(t) = x(t) \cos \omega_s t - y(t) \sin \omega_s t ,$$

where $x(t)$ and $y(t)$ are slowly varying random independent variables which are Gaussian distributed and have second moments equal to N . Since $x(t)$ and $y(t)$ are independent, the joint probability density function is given by the product of the individual density functions.

Hence

$$(63) \quad p(x, y) = p(x) p(y) = \frac{e^{-(x^2 + y^2)/2N}}{2\pi N} .$$

When a signal, $E_s \cos \omega_s t$, is added to the noise, the signal-plus-noise that would exist at the input to the PLL can be represented as

$$(64) \quad v(t) = (x + E_s) \cos \omega_s t - y(t) \sin \omega_s t .$$

The term $(x + E_s)$ represents a Gaussian variable, with E_s as the average value and N the variance. As a result, the sum can be represented by a new variable, $x'(t) = x(t) + E_s$. The random variables x' and y are related to r and θ , the magnitude and phase, respectively, of the envelope by

$$(65) \quad x' = r \cos \theta, \quad y = r \sin \theta .$$

The probability density function can then be expressed as [40]

$$(66) \quad q(r, \theta) dr d\theta = p(x', y) dx' dy = \frac{e^{-[(x' - E_s)^2 + y^2]/2N}}{2\pi N} dx' dy$$

$$= \frac{e^{-E_s^2/2N} r e^{-(r^2 - 2r E_s \cos \theta)/2N}}{2\pi N} dr d\theta .$$

The term $q(r, \theta)$ cannot be expressed as $q(r) q(\theta)$ because the coupling term contains $r \cos \theta$. Hence, r and θ are dependent random variables.

The probability density function of the phase can be found in integrating over the value of the envelope r . Hence,

$$(67) \quad q(\theta) = \frac{e^{-E_s^2/2N}}{2\pi N} \int_0^\infty r e^{-(r^2 - 2E_s r \cos \theta)/2N} dr ,$$

where $s^2 = E_s^2 / 2N$ = signal-to-noise power ratio. The integral can be evaluated by a change of variable and some manipulation[41] to

$$(68) \quad q(\theta) = \frac{e^{-s^2}}{2\pi} + \frac{1}{2} \sqrt{\frac{s^2}{\pi}} \cos \theta e^{-s^2 \sin^2 \theta} [1 + \operatorname{erf}(s \cos \theta)] .$$

For $s^2 = 0$ (no signal) the density function becomes $q(\theta) = \frac{1}{2\pi}$, as expected. For large s , a plot of Eq. (68) is shown in Fig. 25, along

Fig. 25--Probability density function of the phase of a sinusoid plus Gaussian noise.

with the density function for $s = 0$. As can be seen, the phase density function peaks at the signal phase (assumed equal to zero degrees) with a variance proportional to the signal-to-noise ratio. The variance can be expressed as

$$(69) \quad \sigma_{\theta_1}^2 = \frac{1}{2 s^2} = \frac{N}{E_s^2} .$$

The variance of the output of the PLL can now be expressed in terms of Eq. (69), and the closed-loop transfer function by the relationship

$$(70) \quad \sigma_{\theta_z}^2 = \frac{N}{E_s^2} \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \frac{C(j\omega)}{R(j\omega)} \right|^2 d\omega ,$$

which gives an analytical expression for the phase jitter at the output of the PLL. This expression can be integrated analytically for some closed-loop transfer functions, but can be graphically evaluated by use of the Bode plot for any transfer function. Another way of viewing this analysis is that the mean square phase jitter is proportional to the noise bandwidth of the loop, which can be expressed as[42]

$$(71) \quad \text{Noise Bandwidth} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \frac{C(j\omega)}{R(j\omega)} \right|^2 d\omega .$$

CHAPTER V

SIGNAL-TO-NOISE RATIO IMPROVEMENT[43]

A. Introduction

A basic premise of the array approach is that it yields an increase in the signal-to-noise ratio (SNR) of the sum channel output over that of the individual channels comprising the array. This statement must be qualified however because there are several sources of noise that might be encountered in the array channels. First there is the noise that originates within the array, such as that due to amplifiers, mixers, impedance mismatching, cable attenuation, etc. These sources of noise are independent and hence incoherent. A second source of noise can be considered external to the array. This noise can originate from thermal point sources, such as radio stars; from extended sources, such as the sun, milky way, the ground at ambient temperature, etc.; and from noise originating with the signal, such as that which may come from the transmitter. When the array elements are directed at a thermal point source, that source will be correlated in the two array channels, evidenced by the fact that an interferometer pattern can be obtained from two spaced apertures which are directed toward a thermal source. Hence a source of this type will add on a coherent basis in the summing receiver; pattern control, if any, will depend on phase lock loop geometry and i.f. bandwidth.

When speaking of extended thermal sources, the situation becomes more nebulous. Depending on the size of the source compared with the array beamwidth, the source will or will not add coherently. In a situation of this type, the output of an array is the convolution of the array beamwidth with the source distribution. This effect is commonly referred to as antenna smoothing by radio astronomers. Referring to Fig. 10, if the i.f. amplifiers in the phase lock loops were made very narrow, such as with crystal filters, and the loops were operated in Mode I, where the master i.f. signal is used as a reference to the other receivers, the loops would tend to lock on the noise from a thermal point source. The effect of this would be the electronic tracking of the interferometer pattern on the point source. The present i.f. loop bandwidth of 12 Kc prevents this.

Thus when a reference is made to an increase in signal to noise ratio of the sum channel over the individual channels, it is implied that this increase is due to the fact that a portion of the noise existing in the individual channels of the array and most specifically that due to the internal effects will add incoherently in the summing channel. It should be pointed out that the various sources of noise, such as from point and extended thermal sources, would similarly exist in the output of the single antenna which is the aperture equivalent of an array.

Hence when a comparison is desired between an array and the equivalent single aperture, the comparison should mainly consider how the noise levels in the array compare with the noise level obtained from the single antenna, assuming of course equal noise temperatures in the single antenna and in each of the array channels.

In an array of the type being considered here, the SNR will increase by 3 db in the sum channel each time the array size is doubled. Hence a two antenna array will obtain a 3 db increase in SNR over that existing in either channel separately; a four antenna array, a 6 db increase in SNR over the individual channels etc. It is the purpose of this chapter to demonstrate that this increase is actually obtained.

B. Discussion

1. Summing receivers

Assume for simplicity that a two-element phased array is receiving a coherent signal and that the signals are added in phase in a summing circuit. Further assume that the signals at the outputs of the two receivers are not equal, which is a situation often encountered in practice because of unequal signal strengths arriving at two spaced antennas or because of unequal gains in the two channels. Additionally, assume that the noise voltages are not equal at the input to the summing circuit, a situation which again is very often encountered in practice and which is most often caused by unequal noise figures in the two channels.

This situation is shown in Fig. 26, with an amplifier in one arm. In

Fig. 26--Combination of unequal SNR's.

practice, there are amplifiers in both arms of the array; but, again for simplicity, the one amplifier will designate the ratio of the gains in the two channels. The question then arises: What is the gain A for maximum signal-to-noise ratio at the output of the summing circuit ? Since the signals are assumed to be coherent, the SNR at the summer output will be

$$(72) \quad \text{SNR}_{\Sigma} = \frac{(A s_1 + s_2)^2}{A^2 N_1^2 + N_2^2},$$

which is the ratio of the signal power to the noise power.

Letting the SNR in channel 1 be $s_1^2/N_1^2 = \eta_1$, that in channel 2 be $s_2^2/N_2^2 = \eta_2$, and the ratio of the noise powers be $N_2^2/N_1^2 = \beta$,

Eq. (72) can be differentiated with respect to the gain A and the derivative set equal to zero. Carrying out the algebra (see Appendix C), it is found that

$$(73) \quad A = \frac{s_1}{s_2} \frac{N_2^2}{N_1^2} \quad .$$

Thus the gain for maximum SNR is directly proportional to the ratio of the signal strengths and inversely proportional to the noise powers. If this value of gain, Eq. (73) is substituted into Eq. (72) and the algebra (see Appendix D), is carried out, the result is

$$(74) \quad \text{SNR}_{\Sigma} = \frac{s_1^2}{N_1^2} + \frac{s_2^2}{N_2^2} \quad .$$

Thus if the gain in the two channels is properly adjusted, the output SNR is equal to the sum of the SNR's in the two channels; this is the maximum that can be obtained. Many times, particularly in pre-detection combining for tropospheric scatter communications, the signals are weighted into the summer on the basis of signal strength only. This is, of course, an optimum procedure if it can be assured that the noise powers into the summing circuit are equal and remain that way for the duration of the measurement. In many cases this is not an unrealistic assumption. The signal strengths can be conveniently measured by an envelope detector in the IF or AGC line of the receivers, the ratio taken, and the resultant ratio used to control the gain of one of the receivers.

This procedure for optimizing SNR does present problems and may give less than optimum performance when SNR's are low in each

channel, thus encountering thresholding problems in the detectors, or when the noise powers cannot be made equal for a period of time. In these cases, a lower threshold detector is needed, as well as a means for monitoring the noise in each channel during the time a signal is present. Both functions can be accomplished by means of a phase-locked demodulator (shown in Fig. 13). The quadrature detector provides a synchronous detector for AM modulation, while the outputs of the phase-locked loop provide for coherent detection of FM or PM modulation. The threshold performance of such detectors has been treated in the literature[44] and hence will not be discussed here. However, by considering the phase-locked detector as a combination of a synchronous AM detector and a coherent detector for angle modulation, some interesting properties are revealed for noise measurements. From the analysis given in Appendix E, it is seen that the noise power can be measured independently of the signal levels; in this way proper control of the gains in the channels can be obtained.

2. Measurement using RMS voltmeter

The measurements that will be shown were made at the IF output of the summing circuit with an HP 3400A true RMS voltmeter. The individual channel noise and signal-plus-noise measurements were made by disconnecting all receivers from the sum receiver,

except for the one being measured. The summing circuit was adjusted so that it provided equal gains from the four input points to the one output point, and the receiver gains were adjusted so that the noise voltages from the four antennas were equal when measured independently at the summing circuit output. The gains were kept constant during the measurement so that, aside from drift, the only variable was signal strength which was controlled by a signal generator injected into the four antennas. Hence, the optimum signal-to-noise ratio improvement was not obtained, as described in the summing circuit discussion. However, the measured curves will be compared against calculated curves, assuming equal channel gains, so that the signal-to-noise ratio improvement can be clearly seen.

Assume, as in Fig. 27, a four-channel array with equal gains in the four channels and with unequal signal-and-noise voltages in the channels. The signal-plus-noise voltages, as measured at the sum output, due to each channel considered independently are

Fig. 27--Summing of unequal SNR
in a four channel array.

$$\begin{aligned}
 (75) \quad (S+N)_1 &= \sqrt{S_1^2 + N_1^2} && \text{at sum output due to channel 1,} \\
 (S+N)_2 &= \sqrt{S_2^2 + N_2^2} && \text{at sum output due to channel 2,} \\
 (S+N)_3 &= \sqrt{S_3^2 + N_3^2} && \text{at sum output due to channel 3, and} \\
 (S+N)_4 &= \sqrt{S_4^2 + N_4^2} && \text{at sum output due to channel 4,}
 \end{aligned}$$

where the signal and noise voltages are uncorrelated. The independent noise measurements are N_1 , N_2 , N_3 , and N_4 . From the signal-plus-noise and the noise measurements on each channel, the signal-to-noise power ratio can be calculated as

$$(76) \quad \frac{S_n^2}{N_n^2} = \frac{(S + N)_n^2}{N_n^2} - 1 .$$

From this, the signal voltage at the summing circuit output can be found as

$$(77) \quad S_n = N_n \left(\frac{S_n^2}{N_n^2} \right)^{\frac{1}{2}} ,$$

where N_n is measured and S_n^2/N_n^2 is calculated from Eq. (76). Since the signals are coherent, the total signal power at the sum output is

$$\left(\sum_{n=1}^4 S_n \right)^2 .$$

The individual noise powers, being incoherent, combine

so that the total noise power at the sum output is $\sum_{n=1}^4 N_n^2$. Thus the

maximum SNR at the output of the summing circuit, assuming equal channel gains, is

$$(78) \quad \text{SNR}_{\Sigma} = \frac{\left(\sum_{n=1}^4 S_n \right)^2}{\sum_{n=1}^4 N_n^2} .$$

The actual SNR at the sum output is obtained from measurements made when all the channels are connected to the summing circuit. These measurements consist of measuring the total sum-noise power, N_{Σ} , and the total signal-plus-noise power, $(S+N)_{\Sigma}$.

From these measurements, the sum SNR can be calculated as

$$(79) \quad \frac{S_{\Sigma}^2}{N_{\Sigma}^2} = \frac{(S + N)_{\Sigma}^2}{N_{\Sigma}^2} - 1,$$

where all the above measurements are made in terms of RMS voltage.

The discrepancy between Eq. (78) and Eq. (79) is then equal to the deviation of the cohering receivers from theoretical performance.

It should be pointed out that the various time constants in the several loops of the array electronics were adjusted for ease in acquiring the signal and for maintaining lockon during operational conditions. Hence, they possibly do not represent the optimum that could be achieved.

Figures 28 through 32 demonstrate the signal-to-noise ratio improvement obtained with the four array elements. The 3 db S+N/N points represent 0 db SNR, and hence the loops cohere for negative SNR's. The fall-off for higher S+N/N is caused by receiver saturation.

3. Measurement using probability density functions

A second measurement was made of the SNR improvement under dynamic conditions of tracking Echo II. The measurement technique utilized the first probability density functions of the received signals, as explained in Appendix F; the measurements are shown in Fig. 33. The signals received from Echo II were of a pulsed nature with a 28% duty cycle. The first peak in the density

Fig. 28--SNR improvement, Mode I, frequency controlled local oscillator.

Fig. 29--SNR improvement, Mode I, phase controlled local oscillator.

Fig. 30--SNR improvement, Mode II,
frequency controlled local oscillator.

Fig. 31--SNR improvement, Mode II,
phase controlled local oscillator.

Fig. 32--SNR improvement, Mode III,
phase controlled local oscillator.

function is caused by the system noise during the off times of the pulse; the second peak is caused by the presence of signal. The ratio of the two peaks can be related to the SNR, as in Appendix F, with the result that a 3 db improvement was obtained when two of the four antennas in the array were utilized.

4. Measurement to demonstrate
signal coherence and noise
incoherence

A third technique has been used to verify the SNR improvement in the array. This technique involves the measurement of the coherence of the signals in the array and the incoherence of the noise. The coherence of the signals is demonstrated by measuring

Fig. 33--SNR improvement from probability density function.

the phase difference between the crystal oscillator in channel A (Fig. 10) and the voltage-controlled oscillators in channels B, C, and D by means of phase detectors. A typical chart recording of the phase difference is shown in Fig. 34; this phase difference follows that predicted by Eq. (1). This phase difference is the correction that the slave loops must insert on their own channels to maintain their signals in phase with channel A. The sinusoidal characteristic results from a change in the angle θ during the time represented by the chart recording. The change in frequency is a result of the rate of change of phase shift, as explained in Eq. (3). The incoherence of the noise is ascertained by measurements with an RMS voltmeter.

Fig. 34--Signal coherence from 28 Mc phase detector recordings.

CHAPTER VI DATA ANALYSIS AND INTERPRETATION OF ECHO II REFLECTED SIGNALS

A. Echo II description

Echo II is a passive communication satellite which was placed into orbit around the Earth on 25 Jan. 1964. Its purpose is to determine the feasibility of using passive reflectors of rf energy for communication between distant points on the surface of the Earth. The satellite is spherically shaped with a diameter of 135 feet (it is shown in Fig. 35 during a ground inflation test). The satellite is of laminated construction; it consists of a layer of aluminum foil, then a film of mylar, and finally a second layer of aluminum foil. The purpose of this type of construction, as contrasted with the aluminized mylar skin of Echo I, is to provide for rigidity after inflation. After it was placed into orbit the satellite was inflated initially by residual air pressure. Following this, a subliming material was used to provide for additional pressure. The photograph of the initial inflation phase (shown in Fig. 36) was taken from an on-board television camera. Figures 36(a) and (b) show the deployment of the canister containing the folded satellite, Fig. 36(c) shows the separation of the canister, and Figs. 36(d) thru (h) show the development of the inflation from residual air.

Fig. 35--Ground inflation of Echo II.

Fig. 36--Orbital inflation of Echo II.

The satellite was placed into a near polar orbit, with an inclination to the equator of approximately 87° . The mean altitude above the Earth is 600 nautical miles; the period of rotation is 109 minutes; and from a ground station the satellite is visible above the horizon for a maximum of approximately 20 minutes. Generally, the satellite is above the horizon in a group of 3 passes; two such groups occur in a 24-hour period. The first reception of signals made by the phased array described in this paper was on revolution 11; periodic tracking has been maintained since that time.

B. Stationary characteristics of
Echo II data

1. Introduction

The random nature of the signals reflected from Echo II[45, 46] leads to an analysis of the signals on a statistical basis. Once this avenue of analysis is chosen, the concern then becomes the stationary characteristics of the signals; this subject is of importance in determining the number of measurements that must be made on a given amount of data. The purpose of this section is to examine the data in the light of determining the stationary characteristics of the data.

It has been established from measurements made from on-board telemetry transmitters, as well as from other experiments[45], that the satellite is rotating about an axis or spinning in its own orbit.

Because of this complex motion the satellite does not present a fixed aspect angle to an observer. This difficulty is compounded by the fact that signals received from the satellite can be interpreted as if the satellite has a surface roughness and that this roughness may change as a function of time. Hence the analysis of Echo II-reflected signals becomes very complicated in the sense that as measurements are being made on each pass, the elevation and azimuth angles of the tracking antenna are changing; the attitude of the satellite with respect to the observer is changing; and over periods of time, the surface characteristics of the reflector may change. From these observations, it may be seen that a particular Echo II pass can never be exactly duplicated. Interest then is in the question of over what interval the statistical characteristics of the reflected signals remain substantially unchanged.

2. Discussion

In Fig. 37, an ensemble of records[47] is shown; individual units are labelled as $^1f(t)$, $^2f(t)$, etc. One statistic of interest in such an ensemble is the average or mean value of the instantaneous values. Averaging may be done on an ensemble basis at some time, $t = t_1$. The ensemble average at this time is denoted as $\langle f(t_1) \rangle$. Another average can be taken at $t = t_2$, yielding $\langle f(t_2) \rangle$. All possible values of t_1 and t_2 may be assigned, within the limits of the record length. If it is found that, for some range of values of t_1 and

Fig. 37--Time and ensemble averages.

t_2 , the averages are equivalent, i. e., $\langle f(t_1) \rangle = \langle f(t_2) \rangle$, the ensemble is said to be stationary in time. This means that the average is independent of time between t_1 and t_2 , thus implying that only one average need be taken, consequently saving a great deal of labor. If the process is stationary in the interval from t_1 to t_2 , then the average can also be taken in a different way. One member of the ensemble may be chosen, and the average taken over time increments from t_1 to t_2 . As the increment becomes very small, the sum of the instantaneous values becomes the integral and the time average, designated as $\overline{f(t)}$, is formed. The time average cannot be correctly interpreted as the average value of the process unless the ensemble average is independent of time, i. e., stationary.

The ensemble averages of a variable x can be written in terms of the first probability density function, $p(x)$, as

$$(80) \quad m_n = \text{avg } (x^n(t)) = \int_{-\infty}^{\infty} x^n p(x) dx$$

and the time averages can be written as

$$(81) \quad \overline{f^n(t)} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x^n(t) dt$$

In terms of the first probability density function, the ensemble is said to be first-order stationary if $p(x)$ is independent of time, i. e., if the density function measured across the ensemble at $t = t_1$ and at $t = t_2$

are equivalent[48]. With this established, it may then be said that, for a stationary ensemble, each moment (defined by Eq. (80)) has the same value at times t_1 and t_2 , which is a strong result. A more powerful property of the two averages can be established if the time averages and the ensemble averages are equal, i. e.,

$$(82) \quad \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f^n(t) dt = \int_{-\infty}^{\infty} x^n p(x) dx$$

this property is called ergodicity[49], which includes the concept of stationarity; that is, an ergodic process is stationary but the converse is not necessarily true. The ergodic properties of a random process have far-reaching implications, among which is the fact that any measurement of a random process can be confined to a single member of the ensemble if the process is ergodic. By knowing this, the number of computations that need to be made to determine the statistics of a process can be greatly reduced.

When a physical system is involved in a measurement, a time record is normally the form in which the data are recorded and from which the computations are made. For example, the output of the antenna system which receives Echo II signals is recorded on magnetic tape. This tape is then played back into an analog computer which measures various statistical properties of the signal. Alternatively, the tape is played back through a digitizing system which periodically

samples the recorded parameter; this information is then fed to a digital computer which computes the signal statistics. When the statistics are obtained from a time record in this way, some information regarding the stationary characteristics of the data should be known for correct measurement techniques and interpretation.

Strictly speaking, an n member ensemble arises from the output of n identical but independent processes. Obviously this is not feasible, or in many cases possible, to obtain. The problem then is how to obtain the ensemble. Perhaps one way in which this could be accomplished would be to take a number of measurements from a single source or process and arrange the various measurements in an ensemble, as in Fig. 37. A precaution would need to be exercised in doing this: The time between the measurement on one member of the ensemble to the next member must be great enough so that the measurement on the second member would be independent of the measurement of the first member. This time can be obtained from the autocovariance function[50] as the time delay when the autocovariance function becomes zero; this means that there is no correlation between the behavior of the time record for times greater than that found where the autocovariance function equals zero.

In order to obtain reasonable estimates of the probability density functions over the ensemble, a large number of ensemble members is needed. This required large number of members, taken no more frequently than the time for the autocovariance function to become zero, can mean that a very long record is required. The total length of time required may be longer than is possible to make. For example, the Echo II satellite is visible for no longer than 20 minutes, during which time the satellite moves from horizon to horizon and rotates about its axis approximately four times. Hence the use of ensemble records may not always be possible, particularly to obtain stationarity information over appreciable periods of time. Another equivalent approach would be to make use of the time records. Using these, measurements could be made by starting at different times during the time record and continuing for a sufficiently long time to get good convergence to the probability density function. A comparison of these various density functions could result in obtaining a measure of the stationarity of the data and determining the periods of time during which the data are stationary. The density functions obtained in this way can then be compared with the density function obtained from the ensemble measurement; thus a measure of the ergodicity of the process can be obtained.

3. Data

The signals of concern in this report originated from the Passive Satellite Research Terminal at Rome, New York[51], and from Ohio University, Athens, Ohio. They were received at Columbus, Ohio, after reflection from Echo II. The signals were brought to baseband by means of synchronous detectors and then recorded on magnetic tape. Following this, the signals were rerecorded from magnetic tape to a high-speed oscillographic recorder and from these recordings the data were analyzed. The probability density histograms were obtained on a continuous basis by hand calculations, and on a sampled basis by both hand calculations and a digital computer. The autocorrelation function curves were obtained from a digital computer. Because of the large amount of data needed to demonstrate stationary characteristics and ergodic properties, three Echo II passes were chosen. The passes selected were #407, pulsed from Rome, New York; #1730 and #1901 CW from Athens, Ohio. They were chosen as representative of the data received on other Echo II passes.

Since the probability density functions were obtained from limited amounts of data, they are plotted in the form of histograms instead of continuous curves.

The autocorrelation functions shown were taken from data directly obtained from a linear detector. As a result, there is

always an average dc value associated with the data. This results in an autocorrelation function which does not approach zero for large values of time delay. The autocorrelation function is comprised of two parts, the autocovariance function and the mean square values of the average of the waveform. The autocovariance function approaches zero for values of time delay of approximately 0.5 seconds or greater, hence the value of the waveform at increments of 0.5 seconds or greater are independent of each other. This property of the waveform can also be shown by the use of the second probability density functions[52], from which the autocovariance function can be obtained. However, this approach was abandoned because of the much larger number of data points required to obtain the second probability density function.

In an attempt to obtain the statistics of a random waveform, one of the first questions that must be answered is, "how long a sample length is needed to get good convergence to what would be the correct function?" ; or "how many data points are needed?" This question is of importance, particularly in the case of an experiment utilizing Echo II. The shortest possible sample length should be used because of the constantly changing geometry of the reflecting surface. In an attempt to answer this question, the first probability density functions were measured using increasingly long samples of Echo II-reflected signals. The sample lengths started at 5 seconds and were increased

to 60 seconds in consecutive 5-second increments. Figure 38 demonstrates the convergence of the first probability density function for Echo II revolution #407 in 10-second increments from 5 seconds through 50 seconds. The data for this pass were in pulsed form as that it represents a periodic sampling of the fading envelope at 30 times per second. The data were read from the oscillographic records once per pulse at the midpoint of the pulse and represented the average deflection produced by the pulse. In many cases, depending upon the receiving station, recording of the data is not on an instantaneous basis as was done for data received at The Ohio State University. Rather the data are recorded on an integrated basis; hence, the same data used for Fig. 38 was calculated on an envelope basis, that is, an envelope was constructed between the pulses, and the density function was calculated on this basis. These results are shown in Fig. 39. The density functions were also calculated for a CW pass, revolution #1730. Figure 40 shows the convergence as a function of time for the data, considered on a continuous basis, and Figs. 41 and 42 show the same data calculated on a sampled basis with sampling rates of 30 and 50 times a second, respectively.

Fig. 38--Convergence of the first probability density function of pulses for Echo II revolution 407.

Fig. 39--Convergence of the first probability density functions of envelope between pulses for Echo II revolution 407.

(a)

(b)

(c)

(d)

Fig. 40--Convergence of the first probability density functions of Echo II revolution 1730, CW modulation.

(a)

(b)

(c)

(d)

Fig. 41--Convergence of the first probability density functions of Echo II revolution 1730, CW modulation, 30 samples per second.

Fig. 42--Convergence of the first probability density functions of Echo II revolution 1730, CW modulation, 50 samples per second.

By close examination of Figs. 38-42, it can be seen in each case that the density function has achieved a quite reasonable degree of convergence at the end of a 30-second sample. The similarity between Figs. 38 and 39 is good, that is, the density functions of the pulses and of the envelope between pulses, as constructed from pulsed data, are good, as would be expected, since the actual behavior of the envelope between pulses cannot be reconstructed. However, the similarity between Fig. 40 and Figs. 41 and 42 is not nearly so pronounced. Recalling that Fig. 40 comes from CW signals, it could be inferred from the comparison that sampling at 30 or 50 samples per second is not sufficient to accurately reconstruct the received envelope. It should be pointed out here that all of the recorded data contained in this report have been taken with the same system and that the system bandwidth through the detectors and tape recorder was held constant at 1 Kc for all measurements. The bandwidth of the oscillographic recorder was 400 cps.

The same general characteristic is shown in Figs. 43 and 44, which give the density function for revolution #1730 calculated for 30-second increments throughout a 2.5-minute sample. Figure 43 was calculated on a continuous basis and Fig. 44 was calculated from sampled data at 50 times per second. It can also be seen, from Figs. 43 and 44, that the density functions do not remain constant

(a)

(b)

(c)

(d)

Fig. 43--First probability density function for consecutive 30 second time periods for Echo II revolution 1730.

(a)

(b)

(c)

(d)

Fig. 44--First probability density function for consecutive 30 second time periods for Echo II revolution 1730, 50 samples per second.

throughout a 2.5-minute period, signifying that to the extent shown in the figures, the data are not stationary.

Figure 45 shows the equivalent of Fig. 44 but for revolution #1901. Figure 46 demonstrates the density function as obtained from ensemble statistics. The ensemble was obtained from revolution #1730 by dividing a 2.5-minute run of data into 0.5-second samples and arranging them into an ensemble as shown in Fig. 37. A time was then chosen and the value of the function at this time in each of the ensemble members was noted. By selecting the ensemble members to be 0.5-seconds long insures that the value of the function in one ensemble member is independent of the value of the function in the next ensemble member. As noted previously, the 0.5-second value was obtained from the autocovariance function. At this point the dilemma of obtaining the ensemble averages must be faced. If stationarity were to be demonstrated on an ensemble over a period of time, for example, 15 seconds, then each ensemble member should be at least 15 seconds long. However, when this is done, the number of ensemble members is reduced considerably, in fact to the point at which very poor convergence to a density function is obtained. For this reason, the stationary characteristics cannot be shown by use of ensemble averages for Echo data. By comparing Fig. 46,

(a)

(b)

(c)

(d)

Fig. 45--First probability density function for consecutive 30 second time periods for Echo II revolution 1901, 50 samples per second.

Fig. 46--First probability density function from ensemble of records for Echo II revolution 1730.

which was obtained from the ensemble, with Fig. 43, obtained from ensemble members, it may be concluded that the data are not ergodic.

The autocorrelation function for revolution #407, shown in Fig. 47, has been calculated. The autocorrelation function has been computed in 5-second increments from 5 seconds through 30 seconds. The autocorrelation function for revolution #1730 is shown in Fig. 48 calculated in 5-second increments from 5 seconds through 60 seconds. The length of time required for the autocorrelation function to converge is the same length of time required for the second probability density function to converge; the autocorrelation function can be calculated from the second probability density function according to Eq. (83):

$$(83) \quad \phi_{11}(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e_1 e_2 p_{\tau}(e_1, e_2) de_1 de_2 ,$$

where $\phi_{11}(\tau)$ is the autocorrelation function, and e_1 and e_2 are independent random variables describing the amplitude of the curve τ seconds apart. The autocorrelation function can be accurately approximated by an exponential curve, $e^{-K\tau}$, with the value of the exponent K being equal to the initial slope of the autocorrelation function. From this exponential representation, the inverse Fourier transform can be found as

$$e^{-K|\tau|} \longleftrightarrow \frac{K^2}{\omega^2 + K^2} ,$$

Fig. 47--Convergence of autocorrelation function
for Echo II revolution 407.

Fig. 48--Convergence of autocorrelation function
for Echo II revolution 1730.

which yields the power spectral density of the waveform. Thus the value of time required for the initial slope of the autocorrelation function to converge to a steady value gives the length of sample required to accurately compute the power spectral density function. As can be seen in Figs. 47 and 48, this length of time is again on the order of 30 seconds.

Figures 49 and 50 give the autocorrelation functions of revolutions #1730 and 1901 calculated in 30-second increments. The initial slope of the function in each case is very close to being the same, indicating that the power spectral density remains constant for this length of time. The fact that the tails of the autocorrelation functions in Figs. 49 and 50 are not the same is due to the fact that the mean-square value of the average value of the waveform is not the same over the length of time involved, i.e., 2.5 minutes. This is because the data were not corrected for range.

4. Conclusions

From the foregoing, it may be concluded that an optimum time for statistical analysis of Echo II-reflected signals is on the order of 30 seconds which is long enough to converge and short enough to be stationary. Further it may be said that a sampling frequency up to 50 cps is not sufficient to resolve the entire characteristics of the signals. Thus for data taken with a pulsed radar having a PRF of 50

Fig. 49--Autocorrelation function for Echo II revolution
 1730 for consecutive 30 second time periods.

(a)

Fig. 50--Autocorrelation function for Echo II revolution 1901 for consecutive 30 second time periods.

(b)

Fig. 50--Autocorrelation function for Echo II revolution 1901 for consecutive 30 second time periods.

Fig. 50--Autocorrelation function for Echo II revolution 1901 for consecutive 30 second time periods.

or lower, or for computations made on digitized data with a sampling frequency of 50 cps or lower, the insufficiency of the sampling frequency should be considered in interpretations. In regard to stationarity it may be said that the density functions calculated from the data for consecutive 30-second increments do show a change from one increment to the next, but this change is small. The data do not appear to be ergodic, as evidenced by comparing the probability density function, as obtained from the ensemble, with those obtained from ensemble members.

C. Probability density functions of Echo II reflected signals

1. Introduction

Experience in tracking Echo II has shown that on the average the amplitude fading on the received signals is between 5 and 10 db[53] . However, at times fades in excess of 20 db have been noted. These fades can occur at rates from approximately 6 cps up to 250 cps, the maximum observed so far. To demonstrate this, the several oscillographic records shown in Fig. 51 indicate the various types of fading and fading rates that have been observed on approximately 100 of the first 2000 revolutions of Echo II. The records shown in Fig. 51 are of pulsed type signals radiated from the Floyd Passive Satellite Research Terminal in Rome, New York, and received at The Ohio State University, Columbus, Ohio. Also included in Fig. 51

(a)

(b)

Fig. 51--Oscillographic display of received pulses.

(c)

(d)

Fig. 51--Oscillographic display of received pulses.

(e)

(f)

Fig. 51--Oscillographic display of received pulses.

(g)

(h)

Fig. 51--Oscillographic display of received pulses.

is a photographic record of pulses measured at the transmitting site. The received signals appear to have the same character over a range of bistatic angles of from around 50° to 0° . The fading on the received signals can be classified into two types: viz., the pulse-to-pulse fading, called interpulse fading, and the fading within the pulse width (varies from 4 m sec at a minimum range to 10 m sec at maximum range), called intrapulse fading. Intrapulse fading is generally caused by surface roughness of the reflector, while interpulse fading is generally caused by relative motion between the target and the receiver.

A study of the fading characteristics is valuable in many respects. For example, such a study should yield information about the surface of Echo II, which is important in evaluating the Echo II concept, i. e., that of a laminated skin to achieve a rigidized surface after the inflation phase. A second area that could profitably make use of the amplitude fading characteristics of the reflected signals would be in the evaluation and design of a communication system using Echo II. For example, the fading characteristics would be necessary in choosing a type of modulation that would yield maximum information transfer. Also the fading would be needed in the design of certain sections of the receiving terminal equipment.

In this section, the amplitude fading is studied from a statistical point of view, namely, the first probability distribution function of the amplitude of the received signals. The distribution functions have been measured with time as a parameter, i. e., the ordinate of the measured functions is proportional to the amount of time the signals reside within amplitude limits established by the analyzer.

2. Results

A brief discussion of the probability distribution and density functions will be presented before the results are given, in order that they may be more easily interpreted. A random waveform, such as is received from Echo II, may be defined as a time series in which the amplitude is random in nature. Because of this randomness, the properties of the waveform can best be described by its statistical properties. One of these statistical properties which is of particular interest in this paper is the first probability distribution function, which describes the distribution of the amplitudes of the random waveform. Probability will be defined as the ratio of the total amount of time a condition is true to the total time of observation; thus a dimensionless number results. The conditions which are used for rating purposes can be selected in several ways: (a) when the amplitude is greater than a preselected level, (b) when the amplitude is less than a preselected level, and (c) when the amplitude is between

two preselected levels. Condition (c) is the one that will be used throughout this paper. Accordingly (refer to Fig. 52), let the condition

Fig. 52--Coordinate system for probability.

of interest be the presence of the function $f(t)$ within the amplitude range Δx . The amount of time the condition is true is given by τ , where τ is summed over the total time of observation. This process is described by Eq. (84);

$$(84) \quad P[(x + \Delta x) > f(t) > x] = \frac{\sum \tau(x, \Delta x)}{T} = \lim_{T \rightarrow \infty} \frac{\Delta \tau}{T},$$

which reads that the probability that $f(t)$ is less than $x + \Delta x$, but greater than x , is the sum of all τ 's divided by the time of observation T . In some cases, it is more convenient to express probability as a density function. This is accomplished by dividing the probability by the "window" width Δx and letting Δx become

arbitrarily small, thus yielding a function which is no longer dimensionless (Eq. 85).

$$(85) \quad p(x) = \lim_{\Delta x \rightarrow 0} \frac{P[(x + \Delta x) > f(t) > x]}{\Delta x} .$$

The relationship between the probability, sometimes referred to as the probability distribution function, and the probability density function is given by

$$(86) \quad P(x) = \int p(x) dx ,$$

where $P(x)$ is a cumulative function. The properties of $p(x)$ are that $p(x) \geq 0$ and

$$\int_{-\infty}^{\infty} p(x) dx = 1 .$$

Figure 53 portrays several density functions of common deterministic signals along with the density function for Gaussian noise and a sine wave imbedded in Gaussian noise.

Of particular interest here is the 28 per cent duty cycle square wave, which is the waveform transmitted normally by the Rome, New York, terminal. Under actual conditions, the periods of off-time for the pulse contain system (receiving system) noise and the periods of on-time for the pulse contain the signal, the system noise as well as medium noise, i. e., inter- and intra-pulse fading. Thus the density

(a) Sine Wave

(b) Sine Wave
in Gaussian Noise

(c) Gaussian Noise

(d) 50% Duty Cycle
Bi-directional Square
Wave (e) 25% Duty Cycle
Bi-directional Square
Wave (f) 75% Duty Cycle
Bi-directional Square
Wave (g) Repeat of (e) But With
Noise Contamination

Fig. 53--Probability density functions.
of common waveforms.

function will have a spread about maximum and minimum points of the waveform. This is demonstrated in Fig. 53g and is the type of density function which would be obtained if the measurement were made in the receiver IF sections. This measurement is difficult to accomplish because of instrumentation problems. To alleviate these problems, the signals are brought to baseband by means of a linear detector, which modifies the statistics of the signal and hence the density functions. During the pulse off-times, system noise only is present. When modified by the linear detector the system noise will yield the Rayleigh density type function[54]. When the pulse is present, along with system noise and medium noise, the linear detector has varying effects on the statistics of the signal, depending upon the signal-to-noise ratio of the composite waveform. For high SNR, the detector has negligible effect on the density function; while at low SNR, the density function approaches the Rayleigh type[54]. Thus the density function obtained from the pulsed type signals contains two characteristic parts, the noise part and the signal-plus-noise part. As a matter of interest, the density function affords a convenient means for the measurement of signal-to-noise ratio[55]. For CW passes, the density function consists of the signal-plus-noise only, and when the measurement of signal-to-noise ratio is needed for this case, a second measurement is needed to obtain the density function for the noise only.

The first probability density functions for several of the Echo II passes are shown in Fig. 54. The passes were selected to give data just after launch, approximately one month after launch, and approximately four months after launch. In addition, on two of the passes (#235 and 262) a second curve is given for comparison purposes. This curve is the calculated density function for a constant amplitude sinusoidal signal in Gaussian noise with an SNR equal to that of the respective pass on which it is plotted. This is the theoretical curve that would exist if there were no medium noise on the signals. Fading of the signals from Echo II is the same, statistically, as the fluctuations in the amplitude of a sine wave in noise, where the sine wave represents specular (correlated) return and the noise represents scattered (uncorrelated) return[56]. In this way, the ratio of the specular return to the scattered return can be easily found by determining the ratio of mean square sine wave to mean square noise required to give the same variance as the measured probability density function for the signal. The specular-to-scattered ratio is given in Table II for the revolutions of Echo II considered in this report. This ratio is indicative of the roughness of the surface of Echo II. That is, if the ratio is high, one possible inference might be that the surface is smooth enough to give a strong specular component in relationship

(a) Revolution No. 24.

(b) Revolution No. 51.

(c) Revolution No. 163.

Fig. 54--Probability density functions
of selected Echo II revolutions.

(d) Revolution No. 177.

(e) Revolution No. 235

(f) Revolution No. 262.

(g) Revolution No. 1901.

TABLE II
SPECULAR TO SCATTERED POWER RATIO
FOR SELECTED ECHO II REVOLUTIONS

Revolution Number	Specular to Scattered Ratio
24	3.5 db
51	5.2 db
163	6.0 db
177	6.5 db
235	6.5 db
262	9.5 db
1901	9 db

to the scattered component (see Fig. 55). If Echo II is rotating, as all evidence indicates, the scattered components are due to the reflection from small surface irregularities which are constantly changing geometry and hence causing the scattered components to vary in a

Fig. 55--Specular to scattered power ratio.

random fashion in both amplitude and phase, hence the reason for referring to the composite of them as uncorrelated return. The specular component, from the highlight or flare spot on the reflector, should be rather steady in amplitude. It is this line of reasoning which is used to draw correspondence between the specular and scattered components and a sine wave in a random Gaussian noise.

Fitting calculated curves to the probability density functions is but one way in which the specular-to-scattered ratio can be found. Another approach uses the autocorrelation function of the signal and this will be reported. Other investigators[57] have used various other schemes to arrive at this ratio, specifically in regard to surface characteristics from the moon and planets.

It should be pointed out that reference to surface roughness is only qualitative in that it is very difficult to draw a correspondence between surface roughness and the return signal statistics. As a result of this, only the trend in surface characteristics can be inferred during the period of time for which the passes have been analyzed[58].

In all cases it can be seen that the measured curve has more spread than the theoretical curve, signifying the presence of medium noise. A second significant parameter pertaining to the density function is the range of fading (ROF). ROF is defined as the ratio of the power between the levels below which 5 per cent of the return

power is found, and below which 95 per cent of the power is found, commonly referred to as the 5% and 95% points[59]. This parameter is important since it also gives a measure of the surface roughness of the scattering body.

Edison[60] has reported an ROF varying from 3 db for a smooth surface to 19 db for a very rough surface. Rice[61] has predicted that for a purely scattered type return the ROF should be 17.5 db, which is described by the Rayleigh distribution. For Echo II, the ROF as measured is given in Table III for several passes.

TABLE III
RANGE OF FADING FOR SELECTED
ECHO II REVOLUTIONS

Pass Number		Signal Characteristics	Range of Fading
24	Echo II	Pulsed	10.5 db
51		"	9 db
163		"	8.2 db
177		"	7.8 db
235		"	7.0 db
262		"	6.0 db
1901		cw	10.5 db
17042	Echo I	cw	12.5 db

For comparison purposes, measurements were made on Echo I, the first-generation passive satellite, using cw signals from the transmitter at Athens, Ohio. The ROF for Echo I, the density function for which is shown in Fig. 56, was measured at 12.5 db, signifying a surface which is rougher in nature than Echo II. It is again difficult to state quantitatively, however, the specific relationship between the variance of the curves (proportional to the ROF) and a measure of the roughness of the surface.

(a) Not linearized.

(b) Linearized.

Fig. 56--Probability density function for Echo I.

Perhaps the most significant source of error in the measurements is the nonlinearities occurring in the various components comprising the system. Measurement made on the recording system show that if nonlinearities are present they are completely negligible. However, similar measurements made on the receiving system demonstrate that considerable nonlinearity is present. A study of the effect of this nonlinearity on the ROF has shown that an error on the order of 1.5 db low is present in the data coming from the data processing equipment. Because this is a relatively small error, and in view of the lack of knowledge on the correspondence between ROF and surface roughness, it was decided not to perform an error analysis on each of the density functions. Therefore the data presented in Figs. 54 and 56 are the direct output of the data processing equipment. The curves would have a slightly different shape, as demonstrated by the corrected curve in Fig. 56, if the curves were corrected for system nonlinearities.

Another possible source of error in the measurements concerns the length of data sample chosen for analysis. The length of sample chosen has been based on two considerations. The first is that the Echo II aspect angle is constantly changing (as evidenced by data received from on-board telemetry transmitters). Its position with respect to the transmitting and receiving sites is also changing. Secondly, a sufficiently long sample must be chosen so that adequate

convergence of the first probability distribution function can be achieved. Based on these considerations, a sample length of 30 seconds has been chosen for analysis purposes. This aspect of the data from Echo II was discussed in a preceding section.

3. Conclusions

It may be concluded from the foregoing discussion that according to a first probability distribution function analysis of the signals from Echo II and Echo I each is a rough scatterer at 2.27 Gc, with surface roughness increasing according to the order of listing. This conclusion is reached from the range of fading of 6 to 10.5 db and 12.5 db, respectively. The data obtained suggest that the shape of Echo II is generally spherical in nature, as evidenced by a strong specular component, but with a surface roughness, evidenced by the greater variance on the measured data compared with calculated curves based on a sine wave in Gaussian noise. The apparent surface roughness of Echo II seemed to decrease up to revolution #262 and then increased again.

The particular approach to the measurement of surface roughness appears to hold promise but is hampered by a lack of knowledge of a relationship between ROF and a measure of surface roughness.

D. Short-term autocorrelation function[62]

1. Introduction

There are several viewpoints which may be adopted in the analysis of data from a satellite such as Echo II. Because of the fact that the signals have a high fading rate, because the position of the reflector with respect to the observer is constantly changing, and because the orientation of the reflector with respect to the observer is also constantly changing (caused by rotation of the reflector about its own axis), the signals are best handled by statistical methods. Following this line of reasoning, the signals may be handled by the probability distribution functions of the signals as just described, by the power spectrum of the signals[63] or by the autocorrelation function of the signals. This section deals with the autocorrelation function of the signals.

Any one of the three statistical analysis methods will, in effect, act as a check upon the other two. That is, for example, the autocorrelation function can be used to obtain the power spectral density, or vice versa, by the use of Fourier transform methods, and the second probability density function can be used to find the autocorrelation function. However, in addition to the obvious advantage the three viewpoints give in checking the consistency and accuracy of the data, each has distinct advantages for arriving at various parameters contained in the data or for ease in analysis. For example,

the first probability density function affords a direct means of measuring the specular-to-scattered power ratio from a rough reflector; the power spectral density affords a convenient way of measuring the rotation rate of a rough reflector, as well as a qualitative means for comparing surface roughness of rough reflectors; and the autocorrelation function makes possible a quantitative measure of the surface roughness of a rough reflector in terms of the rms surface slope. Additionally, the probability density functions and the autocorrelation functions are best handled by digital computer techniques because of the efficiency in producing numerical comparisons or multiplications; the power spectral density is best handled by analog computer techniques because of the efficiency in making frequency measurements. However, both digital and analog methods can be and are used for obtaining all of the statistical functions of interest.

2. Discussion

In this section, there will be need for the use of the autocorrelation function, defined as

$$(87) \quad \phi(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t) f(t + \tau) dt ,$$

and for the autocovariance function, defined as

$$\begin{aligned}
 (88) \quad \mu_{11}(\tau) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} [f(t) - \bar{f}(t)][f(t+\tau) - \bar{f}(t+\tau)] dt \\
 &= \phi(\tau) - \bar{f}^2(t) .
 \end{aligned}$$

When the average value of the waveform $f(t)$ is zero, then the autocorrelation function equals the autocovariance function. The signals of interest are those of the envelope of the signals reflected from Echo II, as obtained from a linear detector. As a result, the signals can be considered to be composed of a dc value, on which there is superimposed a fluctuating component caused by the fading of the Echo II-reflected signals. It is this waveform that is used in computing the autocorrelation functions, and the autocovariance function is obtained by subtracting from the autocorrelation function the mean square value of the waveform.

a. Specular-to-scattered power ratio

Since Echo II is a rough reflector, the signals scattered from it can be visualized as in Fig. 55. The specular component is a result of reflection from the spherical or spheroidal shape and the scattered components are results of reflection from the various facets caused by the surface roughness of the reflector. Since the reflector is rotating about an axis, the scattered components will have varying amplitudes and phases; hence the composite scattered signals will have somewhat the characteristics of random narrow band Gaussian

noise. It is this line of reasoning that permits one to draw a correspondence between a sine wave in Gaussian noise and the signals scattered from Echo II, where the sine wave represents the specular (correlated) component and the Gaussian noise represents the scattered (uncorrelated) component. The process of interest in this report is shown in Fig. 57, where a sine wave in Gaussian noise is linearly detected and the autocorrelation function taken of the resulting envelope.

The autocorrelation function shown for the signal is typical of the form of the function obtained from Echo II-reflected signals. Also shown is the autocorrelation function of a dc signal and of random noise, the purpose of these being to demonstrate that the Echo II signals fall somewhere between these extremes. The average value at the output of the linear detector is proportional to the specular component of the Echo II-reflected signal, while the rms value of the fluctuations about the average value of the total waveform is proportional to the scattered components of the Echo II-reflected signal. Since the average value of the waveform at the output of the linear detector is uncorrelated with the fluctuating component at the detector output, the autocorrelation function at $\tau = 0$, $\phi(0)$, represents the sum of the mean square value of the fluctuating component and the mean square value of the average value, normalized to unity. For larger values of τ , a point is reached at which the

Fig. 57--Definition of specular to scattered power ratio from the autocorrelation function.

correlation of the fluctuating component with itself approaches zero, and the autocorrelation function at that point is proportional to the mean square value of the average value of the detector output, designated $\overline{f^2}(t)$. Thus the specular-to-scattered power ratio can be found from the autocorrelation function by the relationship (see Fig. 57)

$$(89) \quad \frac{\text{specular power}}{\text{scattered power}} = \frac{\overline{f^2}(t)}{1 - \overline{f^2}(t)} .$$

b. Inverse transforms

By considering only the autocovariance function, it can be established that an exponential function is a very good approximation to the function, particularly for small τ 's (see Section B, this Chapter). This is fortunate indeed in that the inverse Fourier transform may be quickly taken to obtain the power spectral density. This transform is given by

$$(90) \quad e^{-K|\tau|} \longleftrightarrow \frac{K^2}{\omega^2 + K^2} .$$

It should be noted that the -3 db point on the power spectrum occurs at $\omega = K$, and that the power spectrum is Gaussian shaped.

It has been shown that for a rough rotating sphere, contours on the planar view of the sphere contributing to equal doppler shifts consist of straight lines parallel to the axis of rotation of the sphere[64].

Thus regions contributing to the maximum up-doppler or down-doppler frequency shift on a signal returned from the sphere occur from the extremities of the sphere, and the point of zero doppler occurs from the specular point. The maximum up-doppler and down-doppler components are in a one-to-one correspondence with the tails of the power spectral density curves for the signal reflected from the rotating sphere. That is, there cannot be power at a higher frequency in the power spectral density plot than that contributed from the points of maximum doppler shift from the rotating sphere. Thus if the frequency at the tails of the power spectral density curves is taken as the maximum doppler frequency, and knowing the mean radius of the sphere, the rotation rate of the sphere can be computed. Using this fact and data from the power spectral density curves of Echo II, the rotation rate of the Echo II sphere comes out to be approximately 1 revolution in 100 seconds.

This same information can also be obtained from the autocorrelation function in that the frequency at which the power spectral density falls to 5% of its maximum value is the point at which $\omega = 5K$, where K is the constant in the exponential expression for the autocorrelation function, i.e., the initial slope. Thus the equation

$$\begin{aligned}
 (91) \quad f_d &= \frac{2v}{\lambda} = \frac{2\omega r}{\lambda} = \frac{4\pi f r}{\lambda} \\
 &= \frac{5K}{2\pi}
 \end{aligned}$$

can be solved for f , the rotation rate of the sphere in revolutions per second, with the result

$$(92) \quad f = \frac{5K \lambda}{8\pi^2 r} .$$

It is therefore reasonable to assume that for a particular satellite or target, the value of K obtained from the autocorrelation function would be close to a constant value for all measurements. This is, in fact, true; but there should be some variation in the value of K because as when a rough rotating sphere is viewed from a direction perpendicular to the axis of rotation, maximum values of up and down doppler frequency shifts should be observed from the two extremities. As the direction of observation is changed, this doppler should decrease until the observation direction is along the axis of rotation, at which point the doppler frequency caused by reflector rotation should approach zero and the power spectral density should approach an impulse function (slope of zero for the autocorrelation function). This phenomenon has been observed approximately three times in more than 300 Echo II passes which have been tracked. The

phenomenon lasted for approximately 2 minutes in each case. A typical signal strength recording is shown in Fig. 58 for an Echo II pass

Fig. 58--Effect of the coincidence of the axis of rotation and the angle of look.

in which this occurred. Thus the value of K would be expected to be a maximum when the observation direction is orthogonal to the axis of rotation of the reflector, and minimum when the axis of rotation coincides with the direction of observation. Using these facts, the orientation of the axis of rotation of Echo II may be determined by carefully measuring the value of K during consecutive complete passes and noting its variations. This is currently under study and will be reported if some meaningful results can be obtained.

c. Surface roughness

Since the initial observations of the signal fading from the lunar surface (similar in many respects to the signal fading from both Echo II and Echo I), many studies[65-68] have been made to determine the functional relationships between the surface parameters of lunar and planetary surfaces and the signal fading characteristics. Two of these by Daniels[56, 68] have used the autocorrelation function of the envelope of the received signals to draw some inference as to what the surface characteristics of the moon might be, i.e., rms surface slope, surface profile, etc. Taking the work by Daniels at face value, the autocorrelation function of the signal envelope is given by

$$(93) \quad \phi(\tau) = \exp\{-\alpha[1 - \gamma(\tau)]\} \quad ,$$

where $\gamma(\tau)$ is the autocorrelation function of the surface of the reflector corresponding to a fixed value of position, and α is a constant given by

$$(94) \quad \alpha = 16 \pi^2 \frac{\overline{h^2}}{\lambda^2} \quad ,$$

where $\overline{h^2}$ is the mean square surface amplitude. In the lunar case, the value of $\overline{h^2}$ is obtained from optical measurements[68]. The autocorrelation function of the surface in terms of the variable τ is related to the function of distance along the surface of the reflector by the relation $r = v\tau$, where v is the tangential surface velocity

obtained from the known rotation rate and r is the nominal radius of curvature of the reflector.

In the case of Echo II, the surface velocity is known from independent measurements of the on-board telemetry transmitters and from power spectral density measurements, as explained previously. However, the mean square surface amplitude of Echo II is not known, although it is reasonable to assume from the first probability density function measurements that this is less than $(\lambda/4)^2$.

Since the signal envelope autocovariance function is accurately represented by an exponential function,

$$(95) \quad \phi(\tau) = e^{-K|\tau|},$$

solving for the surface autocorrelation function results in

$$(96) \quad \gamma(\tau) = \frac{\alpha - K\tau}{\alpha},$$

where $\gamma(\tau)$ is now a linear function of τ . For typical autocovariance functions of the signal envelope, it may be seen that the autocovariance function becomes very small for values larger than 0.5 seconds, and accuracies are likely to be small if results are projected beyond this point. If this value of τ is translated into distances on the reflector surface (Echo II), it will be seen that $\gamma(\tau)$ represents the surface autocorrelation function in a region approximately 2 feet in

diameter. Hence the portion of $\gamma(\tau)$ deducible from reflected signal measurements is a very small section located near the flare spot on the reflector, and little information other than initial slope can be obtained from $\gamma(\tau)$.

Hayre[69] has shown that many naturally occurring surfaces can be represented by a surface autocovariance function of the form

$$(97) \quad \rho(r) \propto e^{-r/B},$$

where B is the distance at which the autocovariance function drops to $1/e$ of its value at $r = 0$. Hence B is a measure of the horizontal size of the surface perturbations, and, further, $\bar{h}/B$ is a measure of the average slope of the surface. Although the surface autocovariance function $\gamma(\tau)$ is a linear function when computed from the envelope autocovariance function, $\gamma(\tau)$ can be approximated by an exponential function near $\tau = 0$, where it best represents the actual surface autocovariance function. The initial slope of $\gamma(\tau)$ is given by

$$(98) \quad \left. \frac{d \gamma(\tau)}{d \tau} \right|_{\tau=0} = - \frac{K}{\alpha}.$$

The exponential function with this initial slope would then be

$$(99) \quad \gamma(\tau) = e^{-\frac{K}{\alpha} \tau}$$

or

$$(100) \quad \gamma(\tau) = e^{-\frac{K}{\alpha v} \tau} \quad .$$

Matching exponents then with Eq. (97), the value of B can be found as

$$(101) \quad B = \frac{\alpha v}{K} \quad .$$

Thus the average slope would be given by

$$(102) \quad \text{Average surface slope} = \frac{\bar{h}}{B} \quad .$$

d. Discrete frequencies

As mentioned in the introduction of this report, the autocorrelation function of a periodic function is also periodic with the same period. Hence if the autocorrelation function shows some periodic nature, then the envelope of the reflected signal will also have the same periodicity. This fact is displayed in the autocorrelation function of the pulsed passes, where close examination will reveal a periodic component with a 1/6-second period, corresponding to a 6 cps component in the received signal envelope.

As previously noted, there are generally two types of fading which are exhibited in pulsed data recordings. These are interpulse fading, which is a fading in the average value of consecutive pulses, and intrapulse fading, which is fading within the pulse width. The interpulse fading is caused by the rotation of the reflector (Echo II) relative to the observer, while intrapulse fading is caused by the

surface roughness of the reflector. It has been consistently noted in oscillographic recordings of pulsed signals that the intrapulse fading has a 6 cycle component, however this component has not been exhibited in power spectral density recording. At present, this 6 cps component is unexplained.

e. Fading frequencies

There are two measures of the fading which occur on signals returned from rough surfaces. The first of these, referred to as the range of fading (ROF) has been discussed previously (Section C, this Chapter). The standard, using this parameter (ROF), is the surface which gives an ROF of 17.5 db and has been referred to as the Rayleigh surface, which signifies a completely rough surface in terms of the operating wavelength.

The second measure of fading is referred to as the fading frequency[70]. It is defined in terms of the variance spectrum $V(f)$, which is the inverse Fourier transform of the autocovariance function, i.e.,

$$(103) \quad V(f) = \int_{-\infty}^{\infty} \mu_{11}(\tau) e^{j2\pi f\tau} d\tau \quad .$$

In terms of the variance spectrum, the fading frequency is

$$(104) \quad \int_0^{f_e} V(f) df = \int_{f_e}^{\infty} V(f) df \quad ,$$

where f_e is the fading frequency. This statement implies that f_e is the frequency above which and below which equal power resides. Making use of the fact that the autocovariance function is accurately represented by an exponential function, the variance spectrum then becomes

$$(105) \quad V(f) = \frac{K^2}{(2\pi f)^2 + K^2} ,$$

where K is the initial slope of the autocovariance function. Substituting this expression into Eq. (104) results in

$$(106) \quad \arctan \frac{2\pi f_e}{K} = \arctan \infty - \arctan \frac{2\pi f_e}{K} ,$$

which can be solved for f_e with the result that

$$(107) \quad f_e = \frac{0.66 K}{2\pi} .$$

The distance the reflector moves (caused by rotation relative to the observer for Echo II) during one period of the fading frequency is given by

$$(108) \quad d = \frac{v}{f_e} ,$$

where v is the tangential surface velocity of the reflector relative to the satellite. Equation (108) can be reduced further since the approximate rotation rate of Echo II is known. Making this

substitution, Eq. (108) becomes

$$(109) \quad d = \frac{4\pi^2}{K} \quad .$$

Equation (109) should compare rather favorably with Eq. (101), repeated here:

$$(101) \quad B = \frac{\alpha v}{K} \quad .$$

Introducing the appropriate constant α and the known surface velocity, Eq. (101) becomes

$$(110) \quad B = \frac{16\pi^2 \frac{\bar{h}^2}{\lambda^2} \frac{2\pi}{100} \times 67.5}{K}$$

$$= \frac{4\pi^2}{100 \lambda^2 K} \bar{h}^2 \quad .$$

By equating the two equations, i. e., Eqs. (109) and (110), and solving for $\bar{h}^2$, the mean square surface amplitude is given by

$$(111) \quad \bar{h}^2 = \frac{3600}{8\pi \times 67.5} = 2.1 \text{ (inches)}^2$$

or

$$(112) \quad \bar{h} = 1.45 \text{ inches,}$$

which is approximately one-quarter wavelength at the operating frequency. Hence this value may be used in the calculation of the average surface slope in Section c of this Chapter.

E. Long-term autocorrelation function

1. Introduction

Because of the new concept of Echo II (i. e., a rigidized surface), there is considerable interest in ascertaining the shape of Echo II and how this shape varies over a period of time. There are two methods by which this may be done: by optical and by rf techniques. The purpose of this section is to present rf measurements that have been made on the Echo II sphere and to indicate one possible way in which these may be interpreted.

2. Discussion

Because of the random nature of the signals received from Echo II, the data analysis is best handled by statistical means. The analysis that will be used involves the autocorrelation function. The autocorrelation function has been defined in Eq. (87), repeated here:

$$(87) \quad \phi(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t) f(t + \tau) dt ,$$

where $f(t)$ is the envelope of the received signal as obtained from a linear detector. The autocorrelation function consists of two parts, i. e., a contribution from the random part of the waveform and one from the average value. This can be seen by representing the signal envelope as

$$(113) \quad f(t) = f'(t) + \bar{f},$$

where $f'(t)$ is the random waveform which fluctuates about the average value $\bar{f}$. Substituting Eq. (113) into Eq. (87) yields

$$(114) \quad \begin{aligned} \phi(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} f'(t) f'(t + \tau) dt \\ + \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \bar{f}^2 dt. \end{aligned}$$

The first part of Eq. (114) is referred to as the autocovariance function and the second part represents the mean square of the average value of the signal envelope. It has been shown (Section B, this Chapter) that for Echo II-reflected signals, the autocovariance function becomes vanishingly small for values greater than 0.5 seconds (see Figs. 49 and 50). The part of Eq. (114) of interest here is the mean square of the average value, since it is this part which is proportional to the specular reflection of Echo II.

It has been established by means of on-board telemetry transmitters that Echo II is rotating in its orbit with a period of approximately 100 seconds. If the shape of the reflector were not precisely spherical, this rotation would cause a fluctuation in the echo area of Echo II, with the fluctuation being periodic at the rotation angle of the balloon. Because of the mechanics of rotating bodies in a gravitational

field, it may be shown that the axis of rotation would tend to stabilize with respect to inertial space since there is gyroscopic action involved. If this situation does exist, i. e., if the axis of rotation is fixed with respect to inertial space, then during a single pass of Echo II, the aspect angle of the observer with respect to the stabilized axis of rotation can vary over considerable ranges. On certain occasions the aspect angle of the receiving station with respect to the rotation axis has been zero degrees. On these occasions, the fading of the signals, which is caused by the satellite rotation, has ceased for a period of approximately two minutes. However in the largest majority of the passes of Echo II which have been tracked, the aspect angle has been other than zero degrees. If Echo II is spheroidal in shape, the fluctuation in the echo area due to the rotation would be proportional to the aspect angle between the observer and the rotation axis, but the rate of the variations in echo area would be constant and equal to one-half of the rotation rate of the reflector.

The calculation of the long-term autocorrelation functions were performed on an IBM 7094 computer. Data points were obtained by hand from oscillographic recordings made from magnetic tape recordings of the envelope of the received signal. The data points were corrected for detector nonlinearities in the digital computer by means of an auxiliary program before the autocorrelation function analysis was performed.

3. Data

Figures 59 and 60 present the long-term autocorrelation functions of Echo II revolutions #1210 measured on April 26, 1964, and 1901 measured on June 17, 1964. The curves were carried out to time delays of 150 seconds using approximately 7 minutes of data from the total of 10 minutes available. The time delays were made in one-second steps from 1 sec to 150 seconds. The fluctuating curves were those obtained from the computer; the dashed curves were obtained by further smoothing. As can be seen from the dashed curves in both figures, there exists a periodicity which would indicate a reflector rotation rate between 90 and 95 seconds, which is in good agreement with the approximately 100-second period obtained from data from the on-board telemetry transmitters.

By examining the amplitude of the fluctuations of the long-term autocorrelation function, it can be seen that there is a variation on the order of 3%. Since echo area is proportional to the radius of curvature of the reflector and since the circumference of the reflector is that of a 135-foot diameter sphere, it can be calculated that for the aspect angle of the observer, with respect to the rotation axis of Echo II which existed for the times data was taken for Figs. 59 and 60, the ratio of the radii of curvature of Echo II is 1.03. This ratio indicates essentially a spherical reflector.

Fig. 59--Long term autocorrelation function for Echo II revolution 1210.

Fig. 60--Long term autocorrelation function for Echo II revolution 1901.

4. Conclusions

From the foregoing, it may be concluded that Echo II is rotating in its orbit with a period between 90 and 95 seconds and that the ratio of extreme radii of curvature is 1.03, representing essentially a spherical reflector for the aspect angles involved in the measurement.

CHAPTER VII CONCLUSIONS

It has been shown that an array of individually steered, parabolic reflectors can be made the aperture equivalent of a single, larger, parabolic reflector by properly combining the signals from the individual channels of the array. The basis used for this equivalence is the improvement in signal to noise ratio in the sum channel of the array over that available in the individual channels. The improvement in signal-to-noise ratio has been demonstrated under operational conditions of tracking Echo I and II. During this tracking, all of the perturbing influences such as: (a) phase changes due to a changing effective baseline, (b) differential doppler, (c) phase scintillations, (d) feedline effects and (e) equipment phase jitter have been encountered and it has been shown that the phase-locked receivers adaptively control the phase of the individual channels so that signal coherence is maintained in the presence of such perturbances.

In order that close to theoretical performance in signal-to-noise ratio improvement be achieved, it is necessary that the gains of the array channels be made proportional to the signal strength and inversely proportional to the noise power in the respective channels. When this condition is achieved, the signal power to noise power ratio in the sum channel is equal to the sum of the signal power to noise power ratios in the individual channels.

The analysis and design of a phase-locked loop can be made by use of the root-locus and the Bode plot. These methods are used to obtain: (a) a measure of the loop transient response, (b) the characteristics of the error signal, (c) the effect of signal strength on loop bandwidth, (d) the tracking range of the loop and (e) the effects of noise on the loop performance.

The data obtained from tracking Echo II has been analyzed and interpreted in terms of the surface characteristics of the passive reflector. From the first probability density function of the envelope of the received signals, the range of fading and the specular to scattered power ratio was measured. These two parameters indicate that possibly the surface roughness of Echo II improved from the early passes up to approximately revolution #300. From the Gaussian shape of the density function, it might be concluded that the surface roughness is uniformly distributed over the surface of the reflector.

From the short-term autocorrelation function of the envelope of the received signals, the specular to scattered power ratio may again be found and this is in agreement with the value found from the probability density function. Also from the short-term autocorrelation function, the power spectral density can be obtained along with a measure of the surface roughness and the mean slope of the surface roughness.

From the long-term autocorrelation function, the rotation rate of Echo II was measured by the periodicity of the autocorrelation function. This rate was found to be approximately one revolution in 95 seconds, which agrees closely with the value found from measurements made with on-board telemetry transmitters. From the magnitude of the variations in the autocorrelation function, it was concluded that Echo II, at the time of measurement, was very close to spherical.

Finally from a study of the stationary characteristics of Echo II reflected signals, it was concluded that a sample length of 30 seconds was long enough to achieve convergence of the measurements but yet short enough to be stationary.

APPENDIX A COMPOUND INTERFEROMETER[71]

The compound interferometer, which consists of a single large aperture used in conjunction with conventional interferometers, has interesting applications in the field of radio astronomy, as well as applications dealing with coherent signals, such as communication systems and radar astronomy. The advantage of the compound interferometer is that it provides the directivity of the entire aperture, but with the side lobe level of a uniform aperture, i. e., the grating lobes have been suppressed.

The principle of operation of the compound interferometer is the multiplication of several antenna patterns to achieve an approximation to a single-lobe pattern. In addition to this principle, the insertion of a continuous phase shifter into one of the lines of the interferometer allows modulation of the in-phase signal, so that it can be recovered by a synchronous detector. These principles are outlined below to demonstrate the capabilities of the compound interferometer.

Assume that we have a two-element interferometer, as shown in Fig. 61, where $A(\theta)$ and $B(\theta)$ are the one-dimensional voltage patterns of the two elements and θ is the beam angle from end fire. Also assume that a continuously rotating phase shifter is inserted in

Fig. 61--Simple interferometer with phase shifter.

one arm of the interferometer and is effecting an α rad/sec phase shift. The voltage output for this combination is given by

$$(115) \quad E_T = \frac{A(\theta)}{\sqrt{2}} e^{-jk\{d/2\} \cos \theta} + \frac{B(\theta)}{\sqrt{2}} e^{j(k\{d/2\} \cos \theta + \alpha t)}$$

assuming a combiner having equal impedances into the three ports.

The power pattern can then be found as follows:

$$(116) \quad P_T = \frac{1}{2} E_T E_T^* = \frac{1}{2} \left[\frac{A^2(\theta) + B^2(\theta)}{2} + A(\theta) B(\theta) \cos(kd \cos \theta + \alpha t) \right]$$

where $k = 2\pi/\lambda$. The second term in Eq. (116) can be expanded as follows:

$$\begin{aligned}
 (117) \quad & A(\theta) B(\theta) \cos (kd \cos \theta + \alpha t) \\
 & = A(\theta) B(\theta) \cos (kd \cos \theta) \cos \alpha t \\
 & \quad - A(\theta) B(\theta) \sin (kd \cos \theta) \sin \alpha t .
 \end{aligned}$$

If the output of the interferometer system represented by Eq. (117) is used as the input to a synchronous detector, which has for reference a signal derived from the phase shifter drive, i.e., $\cos \alpha t$, the detector will have as an output

$$\begin{aligned}
 (118) \quad \text{Detector}_{\text{out}} & \propto A(\theta) B(\theta) \cos (kd \cos \theta) \\
 & \quad + A(\theta) B(\theta) \cos (kd \cos \theta) \cos 2\alpha t \\
 & \quad + A(\theta) B(\theta) \sin (kd \cos \theta) \sin 2\alpha t .
 \end{aligned}$$

If this output is then filtered with a low-pass filter so that the term containing the second harmonic of the phase-shifting frequency is removed, it can be seen that the output of the system is a dc voltage proportional to the product of the voltage patterns $A(\theta)$ and $B(\theta)$, and of a term representing the interference between the phase center of $A(\theta)$ and the phase center of $B(\theta)$.

Assume now that $A(\theta)$ represents the voltage pattern in the θ plane of a single aperture antenna having a uniform distribution. This aperture can be a circular parabola, cylindrical parabola, or any type of aperture having a uniform amplitude and phase distribution in the θ plane. Thus

$$(119) \quad A(\theta) \propto \frac{\sin\left(\frac{kd_1 \cos \theta}{2}\right)}{\left(\frac{kd_1 \cos \theta}{2}\right)},$$

where d_1 represents the dimension of the aperture in the θ plane and $k = 2\pi/\lambda$ as before. Assume also that $B(\theta)$ represents the voltage pattern of an interferometer system having element separations of d . Starting with the degenerate case of an "interferometer," consisting of a single element located at the edge of the single aperture as shown in Fig. 62(a), and assuming this element to be isotropic, it can be seen that the output of the synchronous detector can be represented by

$$(120) \quad \text{Det}_{\text{out}} \propto \frac{\sin\left(\frac{kd_1 \cos \theta}{2}\right)}{\left(\frac{kd_1 \cos \theta}{2}\right)} \times 1 \times \cos\left(\frac{kd_1 \cos \theta}{2}\right)$$

$$\propto \frac{\sin(kd_1 \cos \theta)}{(kd_1 \cos \theta)}.$$

Thus, the directivity of the single aperture has been increased by two with no resultant increase in the 14 db sidelobe level of the uniform aperture.

With the addition of a second isotropic element at a distance d_1 from the first, as shown in Fig. 62(b), the output of the detector is

Note - The first interferometer element is located immediately adjacent to the single aperture and all element spacings are equal to the length of the single aperture.

Fig. 62--Effects of adding interferometer elements in a compound interferometer.

$$\begin{aligned}
 (121) \quad \text{Det}_{\text{out}} &\propto \frac{\sin\left(\frac{kd_1 \cos \theta}{2}\right)}{\left(\frac{kd_1 \cos \theta}{2}\right)} \cos\left(\frac{kd_1 \cos \theta}{2}\right) \cos(kd_1 \cos \theta) \\
 &\propto \frac{\sin(2kd_1 \cos \theta)}{(2kd_1 \cos \theta)} .
 \end{aligned}$$

Here the directivity has again been doubled and now is four times greater than that of the single aperture, but still with 14 db sidelobes.

This procedure can be continued as shown in Figs. 62(c) and 62(d).

In fact, by induction it can be shown that

$$(122) \quad \text{Det}_{\text{out}} \propto \frac{\sin 2nx}{2nx} \quad n = 1, 2, 3, \dots,$$

where n is the number of interferometer elements.

It also might be pointed out that some control over the type of pattern that can be obtained is available in the three distances d , d_1 , and d_2 . Thus, if the baseline of the simple interferometer is not equal to the length of the single aperture, and if the length of the baseline of the interferometer formed by the single aperture and the simple interferometer is not equal to either of the above, Eq. (121) should be written as

$$\begin{aligned}
 (123) \quad \text{Det}_{\text{out}} &\propto \frac{\sin\left(\frac{kd_1 \cos \theta}{2}\right)}{\left(\frac{kd_1 \cos \theta}{2}\right)} \cos\left(\frac{kd \cos \theta}{2}\right) \\
 &\quad \cos(kd_2 \cos \theta) ,
 \end{aligned}$$

where d_1 = length of single aperture in θ plane,
 d = distance between elements of simple
interferometer, and
 d_2 = distance between phase centers of single
aperture and simple interferometer.

It should be noted here that while the directivity of the compound interferometer can be increased by the addition of more interferometer elements, the gain of the system does not increase in the same proportion. In fact, the gain of the system is in part proportional to the aperture areas of the various apertures comprising the system, assuming of course that the aperture efficiencies, ohmic losses, etc. are also considered.

APPENDIX B WIDE-BAND RECEPTION

Ordinarily, when one considers the response of a two-element interferometer, either analytically or experimentally, a monochromatic source of signal or one contained in a quite narrow band of frequencies is assumed. It is this fact that gives rise to the interference lobes of an interferometer antenna system. In radio astronomy applications, however, practically all noise sources radiate on frequencies which cover a wide band about the center frequency of interest. The energy contained in a narrow sub-interval of the wideband is incoherent, with respect to that in another subinterval, because of the characteristics of thermal noise; but the energy of a single sub-interval appears to be coherent individually to the response of an interferometer. This principle can be used to a certain extent in controlling the response of an interferometer[72], the controlling factor being the reception bandwidth.

The phenomenon of wideband reception can most easily be understood by considering the interferometer as a filter. To do this, the following line of reasoning may be used:

Aperture distribution functions = $E(x/\lambda)$,

Antenna field pattern = $E(\sin \theta)$,

Antenna power pattern = $U(\sin \theta)$, and

Fourier transform pair = $\begin{matrix} \rightarrow \\ \leftarrow \end{matrix}$.

Then it is commonly understood that

$$E(x/\lambda) \begin{matrix} \rightarrow \\ \leftarrow \end{matrix} E(\sin \theta) ,$$

that is, the aperture distribution is the Fourier-Transform of the far field pattern and vice versa[73, 74]. Also, the antenna power pattern is proportional to the square of the voltage pattern, or more properly

$$\begin{aligned} U(\sin \theta) &\propto |E(\sin \theta)|^2 \\ &\propto E(\sin \theta) E^*(\sin \theta) , \end{aligned}$$

where $E^*(\sin \theta)$ is the complex conjugate of the voltage pattern.

Thus one can write

$$U(\sin \theta) \propto \text{FT} [E(x/\lambda)] \cdot \text{FT} [E^*(x/\lambda)] .$$

This expression can most easily be reduced by complex convolution, thus yielding the following result:

$$\text{FT}[E(x/\lambda)] \text{ FT}[E^*(x/\lambda)] \rightleftharpoons \int_{-\infty}^{\infty} E(x/\lambda - \delta) E^*(\delta) d\delta .$$

From the above expression it follows that the antenna power pattern and the autocorrelation of the aperture distribution form a Fourier Transform pair:

$$U(\sin \theta) \rightleftharpoons \int_{-\infty}^{\infty} E(x/\lambda - \delta) E^*(\delta) d\delta .$$

The aperture autocorrelation function can be considered as the spectral frequency response of the antenna, while the power pattern can be considered as the spectral angular response. Some aperture distributions are shown in Fig. 63 along with the associated field pattern and the aperture autocorrelation function.

As can be seen, the frequency response of an interferometer consists of three pass bands located D/λ apart. In the wideband interferometer, the same situation exists except that the location of the two side pass-bands varies with the various frequencies contained in the pass band of the interferometer electronics. However, the center pass band does not move in position but merely changes in width. The resultant frequency response is then represented by the integral of the aperture autocorrelation function over the frequency band of interest. It can be easily seen that when there are two elements in the interferometer, a quite wide frequency band would be

Fig. 63--Examples of far-field patterns and aperture auto-correlation functions for three aperture distributions.

needed in order to effect an improvement in sidelobe levels, and even then, the major effect would be a smearing together of the sidelobes without a great reduction in their level. This is shown qualitatively in Fig. 64, where the smearing of the correlation function of the antenna pattern is caused by the wide bandwidth. Although, this is not the most desirable situation, it would still be quite useable and would have obvious advantages over a conventional interferometer pattern.

This situation however can be improved considerably by increasing the number of elements in the interferometer. This would

Fig. 64--Aperture autocorrelation function and power pattern of a two-element interferometer receiving signals over a wide band of frequencies.

have the following effect on the correlation function for a monochromatic signal source: For simplicity, four elements will be shown (see Fig. 65). Notice there are $(2n-1)$ spikes in the correlation function, where n is the number of elements in the interferometer. Also note that the amplitude of the spikes decreases from the center one. The effect of a wideband signal on the multi-element interferometer correlation function is to fill in the dips between the spikes so that the dotted line is approached in Fig. 65. It should be noted that the form of the correlation function is then approaching that of a uniformly illuminated aperture which yields a sidelobe level of approximately 14 db. Thus, it can be seen from a consideration of the phenomena occurring here that the more elements used and the

Fig. 65--Aperture autocorrelation function and aperture distribution for a four element interferometer.

wider the reception bandwidth, the more closely the operation of a uniform aperture is approximated. It can also be seen that the width of the reception band necessary for destruction of interference is dependent upon the spacing between adjacent elements of the interferometer, since this one factor controls the aperture autocorrelation function for the multi-element case.

This approach should be particularly fruitful for incoherent applications, since one of the goals of the radiometer designer is to obtain as wide an RF bandwidth as is possible. This aids tremendously in reducing the minimum detectable temperature that can be measured with a radiometer receiver.

APPENDIX C
OPTIMUM VALUE OF GAIN IN AN ADAPTIVELY
PHASED TWO-CHANNEL ARRAY TO ACHIEVE
MAXIMUM SIGNAL TO NOISE RATIO

The SNR at the output of the summing circuit in Fig. 26 can be expressed as

$$(124) \quad \text{SNR}_{\Sigma} = \frac{(AS_1 + S_2)^2}{A^2 N_1^2 + N_2^2} \quad .$$

Letting $S_1^2/N_1^2 = \eta_1$, $S_2^2/N_2^2 = \eta_2$ and $N_2^2/N_1^2 = \beta$, Eq. (124) can be reduced to

$$(125) \quad \frac{A^2}{A^2 + \beta} \eta_1 + \frac{2A}{A^2/\sqrt{\beta} + \sqrt{\beta}} \sqrt{\eta_1 \eta_2} + \frac{1}{A^2/\beta + 1} \eta_2 = \text{SNR}_{\Sigma} \quad .$$

Differentiating Eq. (125) with respect to A and setting the results equal to zero, there is obtained

$$(126) \quad \eta_1 \left[\frac{2A}{A^2 + \beta} - \frac{2A^3}{(A^2 + \beta)^2} \right] + \left[\frac{2\sqrt{\beta}}{A^2 + \beta} - \frac{4A^2\sqrt{\beta}}{(A^2 + \beta)^2} \right] \sqrt{\eta_1 \eta_2} - \frac{2A\beta}{(A^2 + \beta)^2} \eta_2 = 0 \quad .$$

By collecting terms and expanding, Eq. (126) can be reduced to

$$(127) \quad A^2 \sqrt{\beta} \sqrt{\eta_1 \eta_2} + A \beta (\eta_1 - \eta_2) + \beta \sqrt{\beta} \sqrt{\eta_1 \eta_2} = 0$$

and solving for A, there is obtained

$$(128) \quad A = \frac{\beta (\eta_1 - \eta_2) \pm \beta (\eta_1 + \eta_2)}{2 \sqrt{\beta} \sqrt{\eta_1 \eta_2}}$$

Choosing the positive sign to obtain a positive gain, there is obtained

$$(129) \quad A = \frac{S_1}{S_2} \frac{N_2^2}{N_1^2} = \frac{S_1 / N_1^2}{S_2 / N_2^2}$$

Thus the optimum value of gain is directly proportional to the signal levels and inversely proportional to the noise powers.

APPENDIX D
MAXIMUM VALUE OF OUTPUT SNR WHEN
THE GAIN SETTING IS OPTIMIZED

If the value of gain found in Eq. (129) is substituted into Eq. (124), there is obtained

$$(130) \quad \text{SNR}_{\Sigma} = \frac{(S_1^2/N_1^2 \quad N_2^2/S_2 \quad + S_2)^2}{S_1^2/N_1^2 \quad N_2^4/S_2^2 \quad + N_2^2} \quad .$$

This equation can be reduced to

$$(131) \quad \text{SNR}_{\Sigma} = \frac{\frac{S_1^4/N_1^4 + S_2^4/N_2^4 + 2 \quad S_1^2/N_1^2 \quad S_2^2/N_2^2}{S_2^2/N_2^4}}{\frac{S_1^2/N_1^2 + S_2^2/N_2^2}{S_2^2/N_2^4}} \quad ,$$

which reduces in final form to

$$(132) \quad \text{SNR}_{\Sigma} = \frac{S_1^2}{N_1^2} + \frac{S_2^2}{N_2^2} \quad .$$

Thus the maximum SNR that can be obtained is the algebraic sum of the signal-to-noise power ratios. By mathematical induction, this can be extended to any number of elements.

APPENDIX E ANALYSIS OF THE PHASE-LOCKED DEMODULATOR FOR MEASURING SIGNAL AND NOISE

The block diagram of the phase-locked demodulator that is being considered in this appendix is shown in Fig. 13. Assume the input to be a carrier-plus-noise as given in Eq. (133)

$$(133) \quad V_{S+N} = A \cos(\omega_c t + \theta_s) + x(t) \cos(\omega_c t + \theta_s) - y(t) \sin(\omega_c t + \theta_s),$$

where $x(t)$ and $y(t)$ are slowly varying random Gaussian variables having zero mean and a mean square value of σ^2 . The output of the AM channel is expressed in Eq. (134).

$$(134) \quad AM_{out} = [A \cos(\omega_c t + \theta_s) + x(t) \cos(\omega_c t + \theta_s) - y(t) \sin(\omega_c t + \theta_s)] \\ B \cos(\omega_c t + \theta_o),$$

which can be reduced to

$$(135) \quad AM_{out} = \frac{AB}{2} [\cos(\theta_s - \theta_o) + \cos(2\omega_c t + \theta_s + \theta_o)] \\ + \frac{B x(t)}{2} [\cos(\theta_s - \theta_o) + \cos(2\omega_c t + \theta_s + \theta_o)] \\ - \frac{B y(t)}{2} [\sin(\theta_s - \theta_o) + \sin(2\omega_c t + \theta_s + \theta_o)].$$

If it is assumed that the loop is locked, i. e., $\theta_s = \theta_o$, and that the output is filtered to remove second harmonic terms, the output can be written as,

$$(136) \quad AM_{out} = \frac{AB}{2} + \frac{B x(t)}{2} ,$$

where B is the peak value of the $VCO/90^\circ$ signal. This signal should be at least four times the p-p value of A so that only the VCO signal controls the gating action of the diodes in the phase detector. The output of the filter following the phase detector thus has two components, i.e., a dc voltage representing the maximum value of the input carrier (modified by the gain of the system) plus a random waveform representing the input noise.

Following the same procedure, the output of FM terminal can be expressed as

$$(137) \quad FM_{out} = \frac{AB}{2} [\sin(2\omega_c t + \theta_s + \theta_o) + \sin(\theta_s - \theta_o)] \\ + \frac{B x(t)}{2} [\sin(2\omega_c t + \theta_s + \theta_o) + \sin(\theta_s - \theta_o)] \\ + \frac{B y(t)}{2} [\cos(\theta_s - \theta_o) - \cos(2\omega_c t + \theta_s + \theta_o)] .$$

When the loop is locked, $\theta_s = \theta_o$ and the harmonics filtered, Eq. (137) can be reduced to

$$(138) \quad FM_{out} = \frac{B y(t)}{2} .$$

Thus the phase-locked loop can be made to act as a narrow-band filter for separating the carrier from the noise. For an unmodulated carrier input plus noise, the AM output represents the carrier as a

dc level plus one-half the random input noise, while the PM output presents one-half the random input noise alone.

Using these properties, it can be seen that by using a narrow-band dc voltmeter at the AM output of the demodulator, the carrier level alone can be measured, and by using an rms meter at the FM output, the noise level alone can be measured, from which the input signal-to-noise ratio (SNR) can be measured.

For AM signals, the signal level can be obtained at the same port as for carrier only, where the dc term is proportional to the carrier power, regardless of whether modulation is present or not. For narrow-band FM, the signal level is obtained by averaging the output of the filter preceding the voltage-controlled oscillator since this filter must be sufficiently wide to pass the modulation frequencies present on the FM signal for proper demodulation. In this case, the output of the AM port will contain the system noise. For wideband FM systems, where both AM and FM would be present on the incoming signal because of the wide deviations present, a portion of the output at both AM and FM ports would be due to the signal level at the input. This situation would have to be carefully studied to separate the signal from the noise but a first-order approach would be to place a controlled limiter in the AM output and adjust it so that limiting on the noise level would not occur but would occur on the signal.

APPENDIX F

MEASUREMENT OF SIGNAL TO NOISE RATIO

1. Introduction

Signal-to-noise ratio involves two terms which are fundamentally different in character. The noise, which is a random waveform, requires different measurement techniques than the signal, which in many cases is a deterministic type waveform. As a result of this basic difference, the signal-to-noise ratio can be defined in several ways, depending upon how the amounts of signal and noise are rated. Because of the random nature of the noise, its rating or measurement is normally based on the mean square value of the noise present. The signal can be measured in several ways, i. e., peak value, average value, mean square value, etc. Hence where signal-to-noise ratio is specified, it must be further clarified. The choice as to how the signal is measured depends upon the particular situation of interest. In many cases, such as in pulse modulation, the value of the peak signal is of primary interest; in other cases, the amount of average or mean square signal could hold more interest.

Ideally the amount of signal present is measured with noise absent and the amount of noise present is measured with the signal absent. In any receiving system, the measurement of the noise is easily accomplished with no signal present; measurement of the

signal with noise absent is difficult or impossible. However the same results are obtained if the signal is measured in the presence of noise so long as the receiving system is linear. This applies specifically to the mixing stages in the receiver, where the local oscillator injection must be sufficient to insure that the statistics of the noise do not change in passing through the mixers. Since the measurement of signal plus noise involves a random waveform, specific type instruments must be used for its measurement, such as those needed for noise alone.

2. Conventional measuring instruments

There are three classes of voltmeters designed for the measurement of random type waveforms. These can be classified as (a) thermoelectric type, (b) true square-law, and (c) synthesized square-law meters. The thermoelectric type meter, such as the HP 3400A, uses bolometers, barreters, or thermocouples which, when properly biased, give accurate indications of the amount of power in the circuit by a change in resistance. True square-law meters utilize a nonlinear element which has an accurate square-law curve, thus giving a reading proportional to the square of the voltage. Such meters employ properly selected diode detectors or utilize the square-law action of certain active elements such as field-effect transistors. The third class of meter, referred to

herein as synthesized square-law, performs the operation of a function synthesizer where a square-law characteristic is formed by the combined action of several nonlinear diodes.

In addition to the above, various types of multiplying circuits are available which form the product between two inputs. These devices can be used in conjunction with a metering circuit to provide the required square-law function by paralleling the inputs of the multiplying circuit. These circuits, while quite accurate, are limited in frequency to a maximum of 50 kc to 100 kc, while the frequency range of some types of RMS voltmeters goes as high as 1000 mc. In most cases, the RMS voltmeters have an adequate frequency response to permit measurements in the IF section of a receiver.

As a matter of interest, the use of average responding meters, which are much more commonly available than true RMS voltmeters, can be used in the measurement of random type waveforms where relative measurements are suitable, if certain precautions are taken. Since the average responding meter does have a meter deflection proportional to the average of the input waveform, and since there is a fixed ratio between the average value and the RMS value for random noise (0.7980), the average responding meter deflection is also proportional to the RMS value of the waveform. The difference between average responding and RMS voltmeters lies in

the allowable crest factor of a waveform it will accurately measure. Crest factor is defined as the maximum ratio of the waveform peak to RMS value. The amplifier stages for an RMS voltmeter are designed to remain linear with a waveform approximately 10 times that required to give full scale deflection. The amplifiers for average responding meters saturate with a waveform approximately 1.5 times that required to give full scale deflection. Since random noise has a high crest factor, care should be exercised to keep the meter deflection at the low end of the scale when random noise is measured with an average responding voltmeter. This is normally accomplished by the use of accurately controllable attenuators at the input to the meter.

3. Random waveform statistics

By applying a knowledge of the statistics of a random waveform to the measurement problem, other techniques can be found that do not require specialized instruments but yet give rather accurate measurements. The statistics of random noise will be discussed here so that the use of an oscilloscope or the probability density function to measure it will be more meaningful.

In referring to random noise, or to a modulated waveform, one can speak of the instantaneous amplitude of the waveform or of the envelope of the waveform. Figure 57(a) which depicts a modulated carrier, demonstrates the distinction between the two

terms. The waveform in Fig. 57(a) could be a modulated waveform or narrow band Gaussian noise. In speaking of the statistics of a random waveform, the statistics could be measured in two possible ways, i. e., time statistics or ensemble statistics (see Chapter VI, Section B). When the statistics from these two methods are equal, the process giving rise to the waveforms are called ergodic. Random noise, which is of concern here, is an ergodic process.

The statistics of primary concern in this appendix are the first probability density functions (see Chapter VI, Section C). As can be seen from Fig. 53, the density function for a sinusoid and noise is a combination of both curves taken individually. This curve and that for the noise alone are of particular concern here as they represent the situation of interest, a carrier in noise (signal-plus-noise) and the noise alone. When these two waveforms are displayed on a properly triggered oscilloscope, the display will be that of the envelope of the waveform (as will be described later). Hence the probability density function of the envelope is needed to properly interpret the oscilloscope presentation.

The probability density function of the envelope of Gaussian noise (as measured at the output of a linear envelope detector) is shown in Fig. 66. This curve is known as the Rayleigh curve, and has the characteristic that the peak of the curve represents the RMS

Fig. 66--Rayleigh density function.

value of the waveform on the amplitude scale. The probability density function of the envelope of a sinusoid plus noise is shown in Fig. 67 as a function of the sinusoid-to-noise ratio (SNR). For low values of the SNR, the curve has the Rayleigh shape as for the noise alone; for high values of the SNR, the curve becomes symmetrical about its peak value and, in fact, approaches a Gaussian curve in the limit.

4. Measurement technique using an oscilloscope

When band-limited Gaussian noise or a sinusoid-plus-noise is displayed on an oscilloscope, and the oscilloscope is triggered by the signal itself, the waveform has the appearance of a sinusoid with a frequency equal to the center frequency of the band-limiting filter

Fig. 67--Probability density function of envelope of sinusoid plus noise.

or of the sinusoid, respectively. An oscilloscope display of these types of signal is given in Fig. 68, where the frequency is 455 Kc. When the oscilloscope is adjusted so that one cycle of the waveform is displayed, the presentation consists of a number of in-phase sinusoids, each of different amplitude, as can be seen by close examination of the photographs in Fig. 68. These sinusoids are a superposition of the individual cycles of the carrier frequency of the waveform in Fig. 57(a). This is to say that because of the triggering action of the oscilloscope, the scope trace can be caused to sweep across the oscilloscope tube at the beginning of each cycle of the carrier frequency. Thus the scope presentation consists of the superposition of an ensemble of waveforms chosen on a cycle-by-cycle basis from the time waveform.

By close examination of the oscilloscope waveforms, it can be seen that there is a varying intensity in the cross section of the waveform taken through its peak. That is, near the base line, the population of sinusoids is small, the population increasing with increasing amplitude and then decreasing with further increases in amplitude. This intensity or population profile through the peak of the waveform is precisely the probability density function of the envelope of the waveform. For the case of noise alone, the intensity profile has the Rayleigh shape, with the point of peak intensity corresponding with

Noise Alone
10 mv/cm

SNR = 1.5
10 mv/cm

SNR = 2.5
10 mv/cm

SNR = 4.7
10 mv/cm

SNR = 5.7
10 mv/cm

SNR = 10.2
20 mv/cm

SNR = 30
50 mv/cm

Fig. 68--Oscilloscope display of sinusoid plus noise.

the RMS value of the noise. Thus by calibrating the vertical scale of the oscilloscope, a rather accurate measure of the noise may be obtained.

The case for a sinusoid-plus-noise is analogous to that for the noise alone, except for the fact that the intensity profile is now the probability density function of a sinusoid-plus-noise, as given by Fig. 67. The point of peak intensity is now dependent upon the signal-to-noise ratio and is, in fact, the sum of the peak signal voltage and the noise voltage. Since the sinusoid and noise voltages are uncorrelated, the voltage at the peak point of intensity is the square root of the sum of the squares of the peak sinusoid voltage and the RMS noise voltage, as given by Eq. (139).

$$(139) \quad V = \sqrt{S^2 + N^2} \quad ,$$

where V is the voltage at the point of peak intensity, S is the square of the peak signal voltage, and N is the RMS noise voltage. Thus by noting the voltage at the point of peak intensity as measured with signal-plus-noise, and knowing the noise voltage as measured with the signal absent, one can solve for the signal alone. Directly solving for S will yield the peak signal voltage and from this, the RMS signal voltage can be solved for. Hence the signal-to-noise ratio can be found on either a peak or RMS signal basis.

An example should serve to illustrate this procedure. Refer to Fig. 68(a), which is the noise alone, and to Fig. 68(d), which is carrier-plus-noise. The point of peak intensity on Fig. 68(a) is at 9 mv, which is the RMS value of the noise alone. The point of peak intensity on Fig. 68(d) is at 29 mv, which represents the sum of the peak signal voltage and the RMS noise voltage. Since the two voltages are uncorrelated, the sum is given by Eq. (139). Solving this equation for S , knowing $V(= 29 \text{ mv})$ and $N(= 9 \text{ mv})$ gives a value of $S = 27.5 \text{ mv}$, which is the peak value of the signal. From this, the RMS value of the signal is 19.5 mv. Thus the ratio of mean signal power to mean noise power is $(19.5/9)^2 = 4.7$. This computation can be reduced to a graphical procedure by using Fig. 69. This graph is entered with the ratio of the signal-plus-noise voltage to the noise voltage and yields the ratio of mean signal power to mean noise power. The second curve on the graph will yield the ratio of peak signal power to mean noise power by entering the graph with the same ratio as above, i.e., the signal-plus-noise voltage to the noise voltage as measured from the oscilloscope presentation. The SNR can similarly be determined by directly measuring the probability density functions.

Fig. 69--Signal to noise ratio versus signal plus noise to noise ratio.

BIBLIOGRAPHY

1. Spilker, J.J., Jr., "Threshold Comparison of Phase-Lock Frequency-Lock and Maximum Likelihood Types of FM Discriminators," WESCON, August 1962.
2. Zolnay, S. L., "A Study of Scintillations at 2 KMC/sec Utilizing Solar Radio Emission," Report 1072-5, 17 August 1962, Antenna Laboratory, The Ohio State University Research Foundation; prepared under Contract AF 30(602)-2166, Rome Air Development Center, Griffiss Air Force Base, New York. AD 295 043
3. von Aulock, W. H., "Properties of Phased Arrays," Proc. IRE, Vol. 48, pp. 1715-1727, 1960.
4. Lerch, C.S., Jr., "Phased Array Radars for Satellite Tracking," Bendix Radio Division, Bendix Corporation.
5. Klass, P.J., "Interest Growing in Phased-Array Radar," Aviation Week, August 20, 1962.
6. Kraus, J. D., Antennas, McGraw-Hill Book Co., Inc., New York, 1950.
7. Brenner, C. H., "Differential Doppler," Report 1072-8, 15 April 1964, Antenna Laboratory, The Ohio State University Research Foundation; prepared under Contract AF 30(602)-2166 with Rome Air Development Center, Griffiss Air Force Base, New York.

8. Zolnay, S. L., op. cit.
9. Eberle, J. W., "The Design of Antenna Arrays for Maximum Signal-to-Noise Ratio," Report 1072-2, 15 March 1961, Antenna Laboratory, The Ohio State University Research Foundation; prepared under Contract AF 30(602)-2166 with Rome Air Development Center, Griffiss Air Force Base, New York.
AD 256 390
10. Caldecott, R., "Some Principles of Low Noise Antenna Design," Report 1041-1, 1 December 1960, Antenna Laboratory, The Ohio State University Research Foundation; prepared under Contract AF 19(604)-6134, Air Force Cambridge Research Center, L. G. Hanscom Field, Bedford, Massachusetts. AD 260 112
11. Copeland, J. R., "Electronic Scanning: A Memorandum on Heterodyne Conversion," Report 903-7, 15 November 1959, Antenna Laboratory, The Ohio State University Research Foundation; prepared under Contract AF 33(616)-6211, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio. AD 233 175
12. Dawirs, H. N., "A Study of the Pattern of the OSU High-Gain Antenna Array," Report 1072-6, June 1962, Antenna Laboratory, The Ohio State University Research Foundation; prepared under Contract AF 30(602)-2166, Rome Air Development Center,

Griffiss Air Force Base, New York. AD 286 337

13. Ksienski, A.A., "Signal Processing Antennas," Microwave Journal, Vol. 4, Pt. I, pp. 77-85, October 1961; Part II, pp. 87-94, November 1961.
14. Waterman, A.T., "A Rapid Beam Swinging Experiment," Trans. IRE, Vol. AP-6, p. 338, 1958.
15. Zolnay, S. L., op. cit.
16. Blair, W.L. and Ammerman, C.R., "Effect of Bandwidth on Receiver Sensitivity," Presented at NAECON, May 12-14, 1958.
17. Brenner, C. H., op. cit.
18. Gangi, A.F., "The Active Adaptive Antenna Array System," Trans. IEEE, Vol. AP-11, pp. 405-414, July 1963.
19. Schrader, J. H., "Receiver System Design for Arraying of Independently Steerable Antennas," Trans. IRE, Vol. SET-8, pp. 148-153, June 1962.
20. Mack, C. L., "Diversity Reception in UHF Long Range Communications," Proc. IRE, 43, pp. 1281-1289, October 1955.
21. Ksienski, A.A., op. cit.
22. Weaver, C. S., "A New Approach to the Linear Design and Analysis of Phase-Locked Loops," IRE Trans. Space Electronics and Telemetry, Vol. SET-5, pp. 166-178, December 1959.

23. Jaffe, R. and E. Rechtin, "Design and Performance of Phase-Lock Circuits Capable of Near-Optimum Performance Over a Wide Range of Input Signal and Noise Levels," IRE Trans. on Information Theory, Vol. IT-1, pp. 66-76, March 1955.
24. Aseltine, J. A., Transform Method in Linear System Analysis, McGraw-Hill Book Co., Inc., New York, 1958.
25. Chestnut, H. A. and R. W. Mayer, Servomechanisms and Regulating System Design, Vol. I, John Wiley and Sons, Inc., New York, 1959.
26. Aseltine, J. A., op. cit.
27. Chestnut, H. A. and R. W. Mayer, loc. cit., p. 412.
28. Chestnut, H. A. and R. W. Mayer, loc. cit., pp. 164-170.
29. Aseltine, J. A., loc. cit., pp. 100-106.
30. Chestnut, H. A. and R. W. Mayer, loc. cit., pp. 383-388.
31. Chestnut, H. A. and R. W. Mayer, loc. cit., pp. 316-351.
32. Evans, W. R., "Instructions for the Spirule," The Spirule Company, Whittier, California.
33. Weaver, C. S., "Thresholds and Tracking Ranges in Phase-Locked Loops," IRE Trans. on Space Electronics and Telemetry, September 1961.

34. Gardner, M.F. and J. L. Barnes, Transients in Linear Systems, Vol. I, John Wiley and Sons, Inc., New York, 1942.
35. Weaver, C. S., op. cit.
36. Chestnut, H.A. and R.W. Mayer, loc. cit., p. 388.
37. Goodman, J.W., R.C. Cumming, B.B. May, I.R. Yaeger, "Some Signal-Suppression Properties of Ideal Limiters," Technical Report 613-2, Systems Techniques Laboratory, Stanford Electronics Laboratories, Stanford University.
38. Weaver, C. S., op. cit.
39. Bennett, W.R., "Methods of Solving Noise Problems," Proc. IRE, Vol. 44, No. 5, pp. 609-638, May 1956.
40. Schwartz, M., Information, Transmission, Modulation and Noise, McGraw-Hill Book Co., Inc., New York, 1959.
41. Schwartz, M., loc. cit., p. 411.
42. George, T.S., "Synchronizing System for Dot Interlaced Color TV," Proc. IRE, February 1951.
43. Zolnay, S. L. and J. W. Eberle, Report in process.
44. Spilker, J. J., op. cit.
45. Zolnay, S. L., "An Analysis of Echo II Reflected Signals at 2 kmc/sec," Report 1072-9, 20 July 1964, Antenna Laboratory, The Ohio State University Research Foundation; prepared under Contract AF 30(602)-2166, Rome Air Development Center,

Griffiss Air Force Base, New York.

46. See Chapter VI, Section C.
47. Lee, C. S., Statistical Theory of Communications, John Wiley and Sons, Inc., New York, 1960.
48. Schwartz, M., loc. cit., pp. 373-381.
49. Lee, C.S., op. cit.
50. Lee, C.S., op. cit.
51. See Chapter VI, Section C.
52. Truxal, J. G., Automatic Feedback Control System Synthesis, McGraw-Hill Book Co., Inc., New York, 1955, p. 430.
53. Zolnay, S. L., op. cit.
54. Schwartz, M., loc. cit., p. 397.
55. See Appendix F.
56. Edison, A.R., R.K. Moore, B.D. Warner, "Radar Terrain Return Measured at Near Vertical Incidence," IRE Trans. on Antennas and Propagation, Vol. AP-8, No. 3, pp. 246-254, May 1960.
57. Brown, W.E., Jr., "A Lunar and Planetary Echo Theory," Journal of Geophysical Research, Vol. 65, No. 10, pp. 3087-3095, October 1960.
58. Hayre, H.S., "Statistical Methods in Radar Astronomy, Determination of Surface Roughness," Journal of Research,

- Vol. 67D, No. 6, pp. 763-765, November-December 1963.
59. Edison, A.R., et al., op. cit.
 60. Edison, A.R., et al., op. cit.
 61. Rice, S.O., "Mathematical Analysis of Random Noise," Bell System Technical Journal, Vol. 23, pp. 282-332, July 1944 and Vol. 24, pp. 46-156, January 1945.
 62. Zolnay, S.L. and J.W. Eberle, Report in process.
 63. Zolnay, S. L., op. cit.
 64. Pettengill, G. H., "Measurements of Lunar Reflectivity Using the Millstone Radar," Proc. IRE, Vol. 48, No. 5, pp. 933-934, May 1960.
 65. Hicks, B.L., et al., "The Spectrum of X-Band Radiation Backscattered from the Sea Surface," Journal of Geophysical Research, Vol. 65, No. 3, pp. 825-839, March 1960.
 66. Daniels, F.B., "Radar Determination of the Root Mean Surface Slope of the Lunar Surface," Journal of Geophysical Research, Vol. 68, No. 2, pp. 449-455, 15 January 1963.
 67. Brown, W.E., "A Lunar and Planetary Echo Theory," Journal of Geophysical Research, Vol. 65, No. 10, pp. 3087-3097, October 1960.
 68. Daniels, F.B., "A Theory of Radar Reflection from the Moon and Planets," Journal of Geophysical Research, Vol. 66, No. 6, June 1961.

69. Hayre, H.S., op. cit.
70. Edison, A.R., "Radar Terrain Return Statistics at Near-Vertical Incidence," Technical Report EE-35, University of New Mexico, October 1960.
71. Covington, A.E. and Broten, N.W., "An Interferometer for Radio Astronomy with a Single-Lobed Radiation Pattern," IRE Trans. on Antennas and Propagation, AP-5, No. 3, pp. 247-255, July 1957.
72. Vitkevich, V.V., "Wide Band Radio Interferometer," Doklady Akad Nauk SSSR, 91, No. 6, 1953.
73. Booker and Clemmov, Proc. IEE, 97, Part 3, p. 11, 1950.
74. van De Hulst, H.E., editor, Radio Astronomy, Cambridge University Press, p. 92, 1957.